# Problem 1

## Part A

**Regression**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 06-FEB-2013 20:06:27 |
| Comments | |  |
| Input | Data | C:\Users\rhaffey\Desktop\CSC 424\assignment2\data\housing.sav |
| Active Dataset | DataSet1 |
| Filter | <none> |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 506 |
| Missing Value Handling | Definition of Missing | User-defined missing values are treated as missing. |
| Cases Used | Statistics are based on cases with no missing values for any variable used. |
| Syntax | | REGRESSION  /MISSING LISTWISE  /STATISTICS COEFF OUTS R ANOVA  /CRITERIA=PIN(.05) POUT(.10)  /NOORIGIN  /DEPENDENT MEDV  /METHOD=ENTER CRIM ZN INDUS CHAS NOX RM AGE DIS RAD TAX PTRATIO B LSTAT. |
| Resources | Processor Time | 00:00:00.03 |
| Elapsed Time | 00:00:00.03 |
| Memory Required | 6828 bytes |
| Additional Memory Required for Residual Plots | 0 bytes |

[DataSet1] C:\Users\rhaffey\Desktop\CSC 424\assignment2\data\housing.sav

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables Entered/Removeda** | | | |
| Model | Variables Entered | Variables Removed | Method |
| 1 | LSTAT, CHAS, B, PTRATIO, ZN, CRIM, RM, INDUS, AGE, RAD, DIS, NOX, TAXb | . | Enter |
| a. Dependent Variable: MEDV | | | |
| b. All requested variables entered. | | | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model Summary** | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .861a | .741 | .734 | 4.74530 |
| a. Predictors: (Constant), LSTAT, CHAS, B, PTRATIO, ZN, CRIM, RM, INDUS, AGE, RAD, DIS, NOX, TAX | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 31637.511 | 13 | 2433.655 | 108.077 | .000b |
| Residual | 11078.785 | 492 | 22.518 |  |  |
| Total | 42716.295 | 505 |  |  |  |
| a. Dependent Variable: MEDV | | | | | | |
| b. Predictors: (Constant), LSTAT, CHAS, B, PTRATIO, ZN, CRIM, RM, INDUS, AGE, RAD, DIS, NOX, TAX | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 36.459 | 5.103 |  | 7.144 | .000 |
| CRIM | -.108 | .033 | -.101 | -3.287 | .001 |
| ZN | .046 | .014 | .118 | 3.382 | .001 |
| INDUS | .021 | .061 | .015 | .334 | .738 |
| CHAS | 2.687 | .862 | .074 | 3.118 | .002 |
| NOX | -17.767 | 3.820 | -.224 | -4.651 | .000 |
| RM | 3.810 | .418 | .291 | 9.116 | .000 |
| AGE | .001 | .013 | .002 | .052 | .958 |
| DIS | -1.476 | .199 | -.338 | -7.398 | .000 |
| RAD | .306 | .066 | .290 | 4.613 | .000 |
| TAX | -.012 | .004 | -.226 | -3.280 | .001 |
| PTRATIO | -.953 | .131 | -.224 | -7.283 | .000 |
| B | .009 | .003 | .092 | 3.467 | .001 |
| LSTAT | -.525 | .051 | -.407 | -10.347 | .000 |
| a. Dependent Variable: MEDV | | | | | | |

## Part B

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables Entered/Removeda** | | | |
| Model | Variables Entered | Variables Removed | Method |
| 1 | LSTAT | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| 2 | RM | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| 3 | PTRATIO | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| 4 | DIS | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| 5 | NOX | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| 6 | CHAS | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| 7 | B | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| 8 | ZN | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| 9 | CRIM | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| 10 | RAD | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| 11 | TAX | . | Forward (Criterion: Probability-of-F-to-enter <= .050) |
| a. Dependent Variable: MEDV | | | |

# Problem 2

**Manova**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 06-FEB-2013 20:34:42 |
| Comments | |  |
| Input | Data | C:\Users\rhaffey\Desktop\CSC 424\assignment2\data\data\_marsh\_cleaned\_homework#2.sav |
| Active Dataset | DataSet1 |
| Filter | <none> |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 165 |
| Syntax | | MANOVA MEHGSWB TURB DOCSWD SRPRSWFB THGFSFC WITH THGSDFC TCSDFB TPRSDFB /  DISCRIM ALL ALPHA(1) /  PRINT SIGNIF(MULTIV UNIV EIGEN DIMENR). |
| Resources | Processor Time | 00:00:00.02 |
| Elapsed Time | 00:00:00.04 |

[DataSet1] C:\Users\rhaffey\Desktop\CSC 424\assignment2\data\data\_marsh\_cleaned\_homework#2.sav

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

The default error term in MANOVA has been changed from WITHIN CELLS to

WITHIN+RESIDUAL. Note that these are the same for all full factorial designs.

\* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* A n a l y s i s o f V a r i a n c e \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*

165 cases accepted.

0 cases rejected because of out-of-range factor values.

0 cases rejected because of missing data.

1 non-empty cell.

1 design will be processed.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

\* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* A n a l y s i s o f V a r i a n c e -- Design 1 \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*

EFFECT .. WITHIN CELLS Regression

Multivariate Tests of Significance (S = 3, M = 1/2, N = 77 1/2)

Test Name Value Approx. F Hypoth. DF Error DF Sig. of F

Pillais .33929 4.05512 15.00 477.00 .000

Hotellings .38686 4.01473 15.00 467.00 .000

Wilks .69630 4.05200 15.00 433.81 .000

Roys .14868

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Eigenvalues and Canonical Correlations

Root No. Eigenvalue Pct. Cum. Pct. Canon Cor. Sq. Cor

1 .17464 45.14311 45.14311 .38558 .14868

2 .13510 34.92338 80.06649 .34500 .11902

3 .07711 19.93351 100.00000 .26757 .07159

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Dimension Reduction Analysis

Roots Wilks L. F Hypoth. DF Error DF Sig. of F

1 TO 3 .69630 4.05200 15.00 433.81 .000

2 TO 3 .81790 4.17630 8.00 316.00 .000

3 TO 3 .92841 4.08707 3.00 159.00 .008

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

EFFECT .. WITHIN CELLS Regression (Cont.)

Univariate F-tests with (3,161) D. F.

Variable Sq. Mul. R Adj. R-sq. Hypoth. MS Error MS F Sig. of F

MEHGSWB .04228 .02443 .47335 .19982 2.36895 .073

TURB .02010 .00184 128.42436 116.66181 1.10083 .351

DOCSWD .13645 .12036 547.75222 64.59251 8.48012 .000

SRPRSWFB .07394 .05668 .00018 .00004 4.28469 .006

THGFSFC .10147 .08473 23331.72485 3849.64750 6.06074 .001

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Raw canonical coefficients for DEPENDENT variables

Function No.

Variable 1 2 3

MEHGSWB .72057 -.61331 .44282

TURB .01490 .00395 .04659

DOCSWD -.12290 -.04565 -.03831

SRPRSWFB -15.97272 77.86417 -98.95910

THGFSFC .00412 -.00985 -.00949

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Standardized canonical coefficients for DEPENDENT variables

Function No.

Variable 1 2 3

MEHGSWB .32611 -.27757 .20041

TURB .16111 .04268 .50364

DOCSWD -1.05314 -.39118 -.32826

SRPRSWFB -.10646 .51898 -.65958

THGFSFC .26750 -.63876 -.61571

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Correlations between DEPENDENT and canonical variables

Function No.

Variable 1 2 3

MEHGSWB -.21383 -.54424 .05581

TURB -.12070 -.03436 .49853

DOCSWD -.89202 -.39006 .02465

SRPRSWFB -.17194 .58138 -.63984

THGFSFC .49143 -.62010 -.52590

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Variance in dependent variables explained by canonical variables

CAN. VAR. Pct Var DEP Cum Pct DEP Pct Var COV Cum Pct COV

1 22.54110 22.54110 3.35130 3.35130

2 23.44119 45.98230 2.79005 6.14136

3 18.76434 64.74664 1.34341 7.48476

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Raw canonical coefficients for COVARIATES

Function No.

COVARIATE 1 2 3

THGSDFC .01142 -.01017 -.01411

TCSDFB -.07756 -.03772 .07279

TPRSDFB -.00297 .00227 -.00422

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Standardized canonical coefficients for COVARIATES

CAN. VAR.

COVARIATE 1 2 3

THGSDFC .73743 -.65693 -.91123

TCSDFB -.95497 -.46446 .89625

TPRSDFB -.55554 .42444 -.79002

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Correlations between COVARIATES and canonical variables

CAN. VAR.

Covariate 1 2 3

THGSDFC -.00951 -.88365 -.46806

TCSDFB -.63909 -.76826 .03666

TPRSDFB -.71407 .14767 -.68433

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Variance in covariates explained by canonical variables

CAN. VAR. Pct Var DEP Cum Pct DEP Pct Var COV Cum Pct COV

1 4.55154 4.55154 30.61395 30.61395

2 5.52607 10.07761 46.42842 77.04237

3 1.64362 11.72123 22.95763 100.00000

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Regression analysis for WITHIN CELLS error term

--- Individual Univariate .9500 confidence intervals

Dependent variable .. MEHGSWB Methyl Mercury in surface water, ng/L

COVARIATE B Beta Std. Err. t-Value Sig. of t Lower -95% CL- Upper

THGSDFC .0003428697 .0489403291 .00073 .47221 .637 -.00109 .00178

TCSDFB .0065911968 .1793289981 .00394 1.67324 .096 -.00119 .01437

TPRSDFB -.0001105168 -.0456875590 .00020 -.56152 .575 -.00050 .00028

Dependent variable .. TURB in situ surface water turbidity

COVARIATE B Beta Std. Err. t-Value Sig. of t Lower -95% CL- Upper

THGSDFC -.0247829117 -.1480850469 .01754 -1.41257 .160 -.05943 .00986

TCSDFB .1488230783 .1695030719 .09518 1.56356 .120 -.03914 .33679

TPRSDFB -.0048860768 -.0845572810 .00476 -1.02741 .306 -.01428 .00451

Dependent variable .. DOCSWD Dissolved Organic Carbon in surface wate

COVARIATE B Beta Std. Err. t-Value Sig. of t Lower -95% CL- Upper

THGSDFC -.0227159190 -.1712436811 .01305 -1.74004 .084 -.04850 .00306

TCSDFB .2761983089 .3968749355 .07082 3.89976 .000 .13633 .41606

TPRSDFB .0058970209 .1287504502 .00354 1.66644 .098 -.00109 .01289

Dependent variable .. SRPRSWFB Soluble Reactive Phosphorus in surface w

COVARIATE B Beta Std. Err. t-Value Sig. of t Lower -95% CL- Upper

THGSDFC -.0000025433 -.0246496088 .00001 -.24187 .809 -.00002 .00002

TCSDFB -.0000992148 -.1832881533 .00006 -1.73916 .084 -.00021 .00001

TPRSDFB .0000091634 .2572148348 .00000 3.21484 .002 .00000 .00001

Dependent variable .. THGFSFC Total Mercury in mosquitofish (Gambusia

COVARIATE B Beta Std. Err. t-Value Sig. of t Lower -95% CL- Upper

THGSDFC .4101116815 .4084975708 .10078 4.06923 .000 .21108 .60914

TCSDFB -1.0939952863 -.2077070328 .54677 -2.00084 .047 -2.17376 -.01423

TPRSDFB -.0294313338 -.0849041902 .02732 -1.07733 .283 -.08338 .02452

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

\* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* A n a l y s i s o f V a r i a n c e -- Design 1 \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*

EFFECT .. CONSTANT

Multivariate Tests of Significance (S = 1, M = 1 1/2, N = 77 1/2)

Test Name Value Exact F Hypoth. DF Error DF Sig. of F

Pillais .30890 14.03484 5.00 157.00 .000

Hotellings .44697 14.03484 5.00 157.00 .000

Wilks .69110 14.03484 5.00 157.00 .000

Roys .30890

Note.. F statistics are exact.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Eigenvalues and Canonical Correlations

Root No. Eigenvalue Pct. Cum. Pct. Canon Cor.

1 .44697 100.00000 100.00000 .55579

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

EFFECT .. CONSTANT (Cont.)

Univariate F-tests with (1,161) D. F.

Variable Hypoth. SS Error SS Hypoth. MS Error MS F Sig. of F

MEHGSWB .12291 32.17032 .12291 .19982 .61510 .434

TURB 177.74614 18782.55136 177.74614 116.66181 1.52360 .219

DOCSWD 1097.07107 10399.39402 1097.07107 64.59251 16.98449 .000

SRPRSWFB .00051 .00675 .00051 .00004 12.13206 .001

THGFSFC 98404.07851 619793.24740 98404.07851 3849.64750 25.56184 .000

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

EFFECT .. CONSTANT (Cont.)

Raw discriminant function coefficients

Function No.

Variable 1

MEHGSWB -.76842

TURB -.00021

DOCSWD .10009

SRPRSWFB 66.13370

THGFSFC .01259

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Standardized discriminant function coefficients

Function No.

Variable 1

MEHGSWB -.34349

TURB -.00226

DOCSWD .80445

SRPRSWFB .42812

THGFSFC .78101

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Estimates of effects for canonical variables

Canonical Variable

Parameter 1

1 2.33781

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Correlations between DEPENDENT and canonical variables

Canonical Variable

Variable 1

MEHGSWB .09245

TURB .14551

DOCSWD .48582

SRPRSWFB .41060

THGFSFC .59600

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

# Problem 3

## Part A

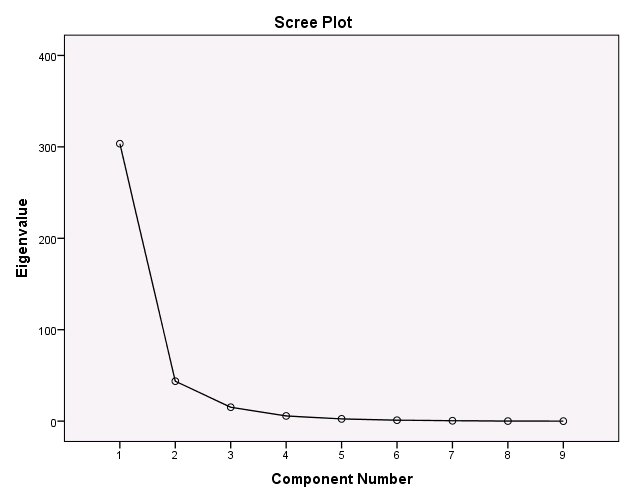
**Factor Analysis**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 06-FEB-2013 20:51:11 |
| Comments | |  |
| Input | Data | C:\Users\rhaffey\Desktop\CSC 424\assignment2\data\country\_data.sav |
| Active Dataset | DataSet1 |
| Filter | <none> |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 26 |
| Missing Value Handling | Definition of Missing | MISSING=EXCLUDE: User-defined missing values are treated as missing. |
| Cases Used | LISTWISE: Statistics are based on cases with no missing values for any variable used. |
| Syntax | | FACTOR  /VARIABLES Agr Min Man PS Con SI Fin SPS TC  /MISSING LISTWISE  /ANALYSIS Agr Min Man PS Con SI Fin SPS TC  /PRINT INITIAL EXTRACTION  /PLOT EIGEN  /CRITERIA MINEIGEN(1) ITERATE(25)  /EXTRACTION PC  /ROTATION NOROTATE  /SAVE REG(ALL)  /METHOD=COVARIANCE. |
| Resources | Processor Time | 00:00:02.27 |
| Elapsed Time | 00:00:04.21 |
| Maximum Memory Required | 11916 (11.637K) bytes |
| Variables Created | FAC1\_3 | Component score 1 |
| FAC2\_3 | Component score 2 |

[DataSet1] C:\Users\rhaffey\Desktop\CSC 424\assignment2\data\country\_data.sav

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Communalities** | | | | |
|  | Raw | | Rescaled | |
| Initial | Extraction | Initial | Extraction |
| Agr | 241.696 | 241.322 | 1.000 | .998 |
| Min | .941 | .374 | 1.000 | .397 |
| Man | 49.109 | 48.260 | 1.000 | .983 |
| PS | .142 | .028 | 1.000 | .195 |
| Con | 2.708 | .954 | 1.000 | .352 |
| SI | 20.933 | 13.565 | 1.000 | .648 |
| Fin | 7.877 | 1.034 | 1.000 | .131 |
| SPS | 46.643 | 40.995 | 1.000 | .879 |
| TC | 1.936 | .629 | 1.000 | .325 |
| Extraction Method: Principal Component Analysis. | | | | |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Total Variance Explained** | | | | | | | |
|  | Component | Initial Eigenvaluesa | | | Extraction Sums of Squared Loadings | | |
|  | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| Raw | 1 | 303.458 | 81.578 | 81.578 | 303.458 | 81.578 | 81.578 |
| 2 | 43.702 | 11.748 | 93.327 | 43.702 | 11.748 | 93.327 |
| 3 | 15.207 | 4.088 | 97.415 |  |  |  |
| 4 | 5.639 | 1.516 | 98.931 |  |  |  |
| 5 | 2.443 | .657 | 99.588 |  |  |  |
| 6 | 1.046 | .281 | 99.869 |  |  |  |
| 7 | .421 | .113 | 99.982 |  |  |  |
| 8 | .065 | .017 | 99.999 |  |  |  |
| 9 | .002 | .001 | 100.000 |  |  |  |
| Rescaled | 1 | 303.458 | 81.578 | 81.578 | 3.352 | 37.244 | 37.244 |
| 2 | 43.702 | 11.748 | 93.327 | 1.557 | 17.303 | 54.546 |
| 3 | 15.207 | 4.088 | 97.415 |  |  |  |
| 4 | 5.639 | 1.516 | 98.931 |  |  |  |
| 5 | 2.443 | .657 | 99.588 |  |  |  |
| 6 | 1.046 | .281 | 99.869 |  |  |  |
| 7 | .421 | .113 | 99.982 |  |  |  |
| 8 | .065 | .017 | 99.999 |  |  |  |
| 9 | .002 | .001 | 100.000 |  |  |  |
| Extraction Method: Principal Component Analysis. | | | | | | | |
| a. When analyzing a covariance matrix, the initial eigenvalues are the same across the raw and rescaled solution. | | | | | | | |



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Component Matrixa** | | | | |
|  | Raw | | Rescaled | |
| Component | | Component | |
| 1 | 2 | 1 | 2 |
| Agr | -15.534 | -.045 | -.999 | -.003 |
| Min | -.033 | .610 | -.035 | .629 |
| Man | 4.726 | 5.092 | .674 | .727 |
| PS | .146 | .079 | .388 | .211 |
| Con | .864 | .456 | .525 | .277 |
| SI | 3.341 | -1.550 | .730 | -.339 |
| Fin | .542 | -.860 | .193 | -.306 |
| SPS | 5.192 | -3.747 | .760 | -.549 |
| TC | .790 | -.065 | .568 | -.047 |
| Extraction Method: Principal Component Analysis. | | | | |
| a. 2 components extracted. | | | | |

## Part B

**Graph**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 06-FEB-2013 21:03:47 |
| Comments | |  |
| Input | Data | C:\Users\rhaffey\Desktop\CSC 424\assignment2\data\country\_data.sav |
| Active Dataset | DataSet1 |
| Filter | <none> |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 26 |
| Syntax | | GRAPH  /SCATTERPLOT(BIVAR)=FAC1\_3 WITH FAC2\_3  /MISSING=LISTWISE. |
| Resources | Processor Time | 00:00:00.45 |
| Elapsed Time | 00:00:00.54 |

[DataSet1] C:\Users\rhaffey\Desktop\CSC 424\assignment2\data\country\_data.sav

