# The Hong Kong Polytechnic University 

# MM604 Statistical Analysis for Management 

## Research

## Assignment 1

Thomas Wu (02715615G)

1. American manufacturers are concerned with the increase of Japanese produced television sets in the US retail markets after an import quota was lifted. Inventory levels of imported television sets are compared one month before and one month after the lifting of the import quota to determine if there is an significant increase. A paired ttest can be used to compare the before and after inventory levels. The hypothesis to be tested are:
$\mathrm{H}_{0}$ : Average inventory levels of imported TVs are the same before and after the lifting of the import quota.
$H_{1}$ : Average inventory levels of imported TVs are NOT the same before and after the lifting of the import quota.

T-Test
Paired Samples Statistics

|  |  |  |  | Std. Error <br> Mean |  |  |
| :--- | :--- | :---: | ---: | ---: | ---: | :---: |
| Pair 1 | AFTER | 149.75 |  | 8 | 57.375 | 20.285 |
|  | BEFOR | 137.38 |  | 8 | 55.926 | 19.773 |

Paired Samples Correlations

|  |  | N | Correlation | Sig. |
| :--- | :--- | ---: | ---: | ---: |
| Pair 1 |  <br> BEFORE |  | 8 | .946 |

Paired Samples Test

|  |  | Paired Differences |  |  |  |  |  | df | Sig. (2-tailed) |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Std. Error95\% Confidence Interval <br> of the Difference |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| Pair 1 | AFTER BEFORE | 12.38 | 18.677 | 6.603 | -3.24 | 27.99 | 1.874 |  | 7 | . 103 |

The two-tailed paired t-test indicates that the difference between the before and after inventory level is significant at the $10.3 \%$ level. Since the study is only concerned with an increase in the level of inventory, a one-tailed paired ttest should be used. To adjust for this, the significant level is halved to $5.15 \%$.

Based on the actual significant level of $5.15 \%$ and the test level of $5 \%$, we cannot reject the null hypothesis that the mean level of inventory is the same before and after the lifting of the import quota. Hence we conclude that the level of inventory of imported television sets has not statistically increased significantly at a 5\% level after the lifting of the import quota.

The histogram below graphs the distribution of the change in inventory level. It appears that the difference might not come from a normal distribution. In addition, the sample size of 8 is small. As a result, a non-parametric test (Wilcoxon Test) is used to confirm the earlier paired t-test results.


DIFF

The non-parametric Wilcoxon Signed Ranks test results below indicate that the difference in inventory level is significant at $6.8 \%$ level. As before, since we are interested in the increase of inventory level only, the actual significant level of $6.8 \%$ should be halved to $3.4 \%$ to determine the significant level for a one-tailed test. Based
on this result, we can reject the null hypothesis that the mean level of inventory is the same before and after the lifting of the import quota. As a result, we conclude based on the non-parametric Wilcoxon test that the level of inventory of Japanese produced television sets has statistically increased significantly at a $5 \%$ level after the lifting of the import quota.

Wilcoxon Signed Ranks Test
Ranks

|  |  | N | Mean Rank | Sum of Ranks |
| :--- | :--- | ---: | ---: | ---: |
| BEFORE - AFTER | Negative Ranks | 6(a) | 5.17 | 31.00 |
|  | Positive Ranks | 2(b) | 2.50 | 5.00 |
|  | Ties | 0(c) |  |  |
|  | Total | 8 |  |  |
|  |  |  |  |  |

Test Statistics(b)

|  | BEFORE - <br> AFTER |
| :--- | ---: |
| Z | $-1.825(\mathrm{a})$ |
| Asymp. Sig. (2-tailed) | .068 |

a Based on positive ranks.
b Wilcoxon Signed Ranks Test

Since we obtained contradictory conclusions using parametric and non-parametric tests, a firm statistical conclusion cannot be drawn. To rectify this situation, an increase in sample size should be obtained and both the tests should be performed again using the expanded sample set In additions to the small sample size, there are several other shortcomings of the current data set which are described below. In order to effectively present their view points, the American manufacturers should address these issues before further pressing their case for a reincarnation of the import quota.

Caution should be used in the interpretation of the results of this statistical analysis as the data set has several shortcomings. First, the change in inventory level between the two time periods could be due to reasons other than then lifting of the quota, meaning that the removal of the import quota might not be the cause of the inventory increase. The retailers might increase the general level of all their inventories due to seasonal sales reasons; as such, the number of US made television sets will experience the same increase. Second, the statistical analysis only provides information on the relationship between the pair of numbers. The analysis does not imply causal relationship. Hence, the results of the statistical analysis does not provide evidence that the lifting of the quota causes the increase in inventory of imported TV sets. Third, the data does not take into account the alternative choice of non-Japanese made television sets. To make their case more persuasive, the American manufacturers should use the ratio of Japanese made to American made television sets as the basis of the study for the increase. Fourth, the use of 8 retail outlets might be too small to be representative of the population. In particular, the origins of television sets being sold are affected by local customer profiles and demographics. Fifth, increase in inventory of Japanese produced television set does not mean that they are dumping. The increase might be due to higher demand for their products because of higher quality. Sixth, dumping is defined as selling below cost. Since cost factors are not included in this study, it would be inappropriate to draw any conclusion on cost or dumping based on the results on this analysis. Seventh, the survey collected was based on imported television sets which could include television sets made in Korea, Germany or other non-Japanese countries.
2. In this study of leadership style and task performance, there are two independent variables, leadership style and nature of task. It is hypothesized that task performance are affected by both the style of the leadership and the nature of the task. Ten groups of soldiers were used and task performance scores are obtained for two styles of leadership and two types of task.
$\mathrm{H}_{0}$ : Task performance is not affected by the leadership style and the nature of task.
$H_{1}$ : Task performance is affected by the leadership style and the nature of task.
The bar chart below summarizes the finding.


A preliminary inspection of the bar graph indicates that there appears to be a difference in performance dependent on the style of leadership and nature of task. In particular, it appears that authoritarian leadership style is better suited for structured tasks like gun assembly; on the other hand, democratic leadership style is better suited for group problem solving tasks.

The box-and-whiskers plot below indicates that the distribution of the task performance does not overlap and there is difference in task performance between the leadership style and nature of task.


The hypothesis to be tested is that leadership style does not affect the task performance of structured task (i.e. gun assembly) and group problem (i.e. problem solving), meaning that the performance scores are not significantly different between the two types of tasks. Since the performance scores are marks awarded, we are trying to compare the equality of the population means using leadership style and nature of task as classification criteria. As such, two-way ANOVA is used for the analysis.

Univariate Analysis of Variance

Descriptive Statistics
Dependent Variable: SCORE

| CRIT | LEAD | Mean | Std. Deviation | N |
| :--- | :--- | :--- | :--- | :--- |
| 1.00 | 1.00 | 15.2000 | 2.28035 | 5 |
|  | 2.00 | 7.8000 | 2.77489 | 5 |
|  | Total | 11.5000 | 4.57651 | 10 |
| 2.00 | 1.00 | 7.0000 | 3.80789 | 5 |
|  | 2.00 | 15.8000 | 2.58844 | 5 |
|  | Total | 11.4000 | 5.56177 | 10 |
| Total | 1.00 | 11.1000 | 5.23768 | 10 |
|  | 2.00 | 11.8000 | 4.91709 | 10 |
|  | Total | 11.4500 | 4.95745 | 20 |

Levene's Test of Equality of Error Variances(a)
Dependent Variable: SCORE

| $F$ | df1 | df2 | Sig. |
| :--- | :--- | :--- | :--- |
| .400 | 3 | 16 | .755 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+CRIT+LEAD+CRIT * LEAD

Tests of Between-Subjects Effects
Dependent Variable: SCORE

| Source | Type III Sum <br> of Squares | df | Mean Square | F | Sig. |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Corrected Model | $330.550(a)$ | 3 | 110.183 | 12.925 | .000 |
| Intercept | 2622.050 | 1 | 2622.050 | 307.572 | .000 |
| CRIT | .050 | 1 | .050 | .006 | .940 |
| LEAD | 2.450 | 1 | 2.450 | .287 | .599 |
| CRIT * LEAD | 328.050 | 1 | 328.050 | 38.481 | .000 |
| Error | 136.400 | 16 | 8.525 |  |  |
| Total | 3089.000 | 20 |  |  |  |
| Corrected Total | 466.950 | 19 |  |  |  |

a R Squared $=.708$ (Adjusted R Squared $=.653$ )

The Levene's test of equality of error variances cannot reject the null hypothesis that the error variance is equal across the four groups of observations.

The result of the two-way ANOVA rejects the null hypothesis that the means of the groups are the same. The "Corrected Model" is significant at the $0.015 \%$ level and the interaction between leadership style and nature of task ("CRIT*LEAD") is significant at $0.001 \%$ level. The "Corrected Model" result rejects the null hypothesis that the
population means are the same across the four groups of observations. The "CRIT*LEAD" result indicates that there are significant interaction between these two classifications. The plot below also confirms the indication that there are interaction between the leadership style and the nature of task.

## Estimated Marginal Means of Performance Scort



Leadership Style

Based on the analysis of the results of the study, we found that leadership style does affect the task performance dependent on the nature of the task. That is, for structured tasks like gun assembly, an authoritarian leadership style will produce better performance; for group problem solving, a democratic leadership style will produce better performance.

## 3. ABC Shoes Company

ABC Shoe Company wants to measure sales potential against four independent variables. These four independent variables are (a) percentage of US footwear sales $\left(\mathrm{X}_{1}\right)$, (b) number of accounts services $\left(\mathrm{X}_{2}\right)$, (c) percentage of US personal income $\left(\mathrm{X}_{3}\right)$, and (d) percentage of non-white population $\left(\mathrm{X}_{4}\right)$.

The basic information on the territory sales (Y, in thousand dollars) and the four independent variables are provided below. Based on the review of the histograms, it appears that all the variables have distributions that are approximately normal. Scatterplot matrix is prepared to confirm linear relationships between dependent and independent variables. Independent variable $X_{4}$ was transformed by taking the natural log as there appears to be a non-linear relationship with the dependent variable prior to the transformation.

Description of Variables.


[^0]Histograms of the variables appear to have distributions that are approximately normal.


Scatterplot matrix using the original values of the variables.

| Percentage of US Fop |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Number of Accounts |  |  |  |
|  |  | Percentage of US Per |  |  |
|  |  |  | Percentage of Non-V |  |
|  |  |  |  $\square$  <br>  $\square$  <br>  $\square$  <br> $\square$ $\square$  <br> $\square \square$ $\square \square$ $\square$ <br> $\square$ $\square$  | Sales |

Scatterplot matrix with $\mathrm{X}_{4}$ transformed by taking its natural log.

| Percentage of US Foo |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Number of Accounts |  |  |  |
|  |  | Percentage of US Per |  |  |
|  |  |  | In (\% of Non-White P |  |
|  |  |  |  | Sales |

(a) Using SPSS, the regression equation is:

Sales (in thousand dollars) $=\mathbf{- 2 1 8 . 0 5 + 1 2 9 . 6 7}$ (\% of US footwear sales)
The regression is significant at the $0.001(0.1 \%)$ level. A stepwise regression is used and independent variables $X_{2}, X_{3}$ and $X_{4}$ are found not to be significant and hence excluded from the regression equation.

Coefficients(a)

a Dependent Variable: Y Sales
(b) About $70 \%$ of the variation in sales among the territories is explained by the regression equation in (a) above based on the adjusted $R^{2}$ of 0.698 .

Model Summary(b)

| Model | R | R Square | Adjusted R <br> Square | Std. Error of <br> the Estimate | Durbin- <br> Watson |
| :--- | :--- | ---: | ---: | ---: | ---: |
| 1 | .853 (a) | .728 | .698 | 209.2867 | 1.157 |

[^1](c) Explanatory variable $\mathrm{X}_{1}$ of percentage of US footwear sales is a significant variable at the $1 \%$ level as shown in the table in part (a).

The other variables are not significant at the $1 \%$ level as shown in the table below of excluded variables. Variable $X_{2}, X_{3}$, and $X_{4}$ are only significant at the $16 \%, 14 \%$ and $74 \%$ level respectively.

Excluded Variables (b)

| Model |  | Beta In | t | Sig. | Partial Correlation | Collinearity Statistics |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  | Tolerance | VIF | Minimum Tolerance |
| 1 | X2 Number of Accounts Services X3 | .387(a) | 1.567 | . 156 | . 485 | . 427 | 2.342 | . 427 |
|  | Percentage of US Personal Income | -.320(a) | -1.663 | . 135 | -. 507 | . 681 | 1.468 | . 681 |
|  | X5 In (\% of Non-White Pop) | -.067(a) | -. 354 | . 733 | -. 124 | . 942 | 1.062 | . 942 |

a Predictors in the Model: (Constant), X1 Percentage of US Footware
b Dependent Variable: Y Sales
(d) Using the regression equation in (a), an expected Y is computed for each territory in the table below.

```
a= -218.0541
b = 129.672
```

| Territory | $\mathrm{X}_{1}$ | Actual Y | Expected Y | Deviations |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |
| 1 | 8.6 | $1,352.0$ | 897.1 | 454.9 |
| 2 | 5.0 | 336.0 | 430.3 | $(94.3)$ |
| 3 | 7.6 | 682.0 | 767.5 | $(85.5)$ |
| 4 | 2.1 | 147.0 | 54.3 | 92.7 |
| 5 | 3.5 | 378.0 | 235.8 | 142.2 |
| 6 | 7.0 | 732.0 | 689.6 | 42.4 |
| 7 | 8.8 | $1,031.0$ | 923.1 | 107.9 |
| 8 | 6.3 | 498.0 | 598.9 | $(100.9)$ |
| 9 | 10.2 | 826.0 | $1,104.6$ | $(278.6)$ |
| 10 | 5.3 | 357.0 | 469.2 | $(112.2)$ |
| 11 | 3.9 | 119.0 | 287.7 | $(168.7)$ |

The scatterplot below graphs the relationship between territory sales and the percentage of US footwear sales $\left(\mathrm{X}_{1}\right)$. While most of the actual sales are relatively close to the regression line, territory 1 is almost an outlier that is way above the expected sales. Territory 9 is also quite far below the regression line.

(e) Most of the actual observed sales are relatively close to the expected sales as predicted by the regression equation, as can be seen from the scatterplot in (d) above. The red line is the regression line, and most of the observed territory sales are slightly above or below the regression line in a random manner.

Observed sales of Territory 1 and Territory 9 appear to be farther away from the expected sales based on the regression equation. Territory 1 has about the same level of percentage of US footwear sales as Territory 7, but its observed sales is way above
the expected amount. Territory 9 has the highest level of percentage of US footwear sales, but its actual sales is drastically below expected sales based on the regression equation.
(f) We have performed a stepwise regression, and the only independent variable remaining in the regression equation is $X_{1}$, the percentage of US footwear sales of the territory. Based on the regression equation from (a) above, each percent change in the percentage of US footwear sales in the territory is positively related to a change in sales of $\$ 129,672$. The other variables, $X_{2}, X_{3}$ and $X_{4}$, are found to be insignificant even at 10\% level (see (b) above).

The table below shows the result of the regression analysis that has maintained all the variables in the analysis. This analysis is performed to ensure that variables $X_{2}, X_{3}$ and $X_{4}$ are not excluded incorrectly.

| Coefficients ${ }^{\text {a }}$ |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Model |  | Unstandardized Coefficients |  | $\begin{gathered} \hline \begin{array}{c} \text { Standardized } \\ \text { Coefficients } \end{array} \\ \hline \text { Beta } \\ \hline \end{gathered}$ | t | Sig. | 95\% Confidence Interval for B |  | Collinearity Statistics |  |
|  |  | B | Std. Error |  |  |  | Lower Bound | Upper Bound | Tolerance | VIF |
| 1 | (Constant) | -115.177 | 253.197 |  | -. 455 | . 665 | -734.727 | 504.373 |  |  |
|  | X1 Percentage of US Footwear | 107.905 | 38.228 | . 710 | 2.823 | . 030 | 14.365 | 201.445 | . 338 | 2.954 |
|  | X2 Number of Accounts Services | 3.209 | 1.757 | . 439 | 1.826 | . 118 | -1.091 | 7.508 | . 371 | 2.695 |
|  | X3 Percentage of US <br> Personal Income | -38.326 | 24.342 | -. 307 | -1.574 | . 166 | -97.888 | 21.236 | . 562 | 1.779 |
|  | X5 In (\% of Non-White Pop) | -43.148 | 120.592 | -. 063 | -. 358 | . 733 | -338.226 | 251.929 | . 685 | 1.459 |

a. Dependent Variable: Y Sales

The regression equation based on this analysis is:
Sales $=-115.177+107.905 X_{1}+3.209 X_{2}-38.326 X_{3}-43.148 X_{4}$

Based on this analysis and ignoring the level of significance of the independent variables, the effects of the independent variables are as follow:

|  | Independent Variables | Coefficients | Impact on Dependent Variable |
| :---: | :--- | :---: | :--- |
| $\mathrm{X}_{1}$ | Percentage of US <br> footwear sales | 107.905 | A percentage increase in US footwear sales <br> in the territory will increase expected sales <br> by $\$ 107,905$. |
| $\mathrm{X}_{2}$ | Number of accounts <br> services | 3.209 | A percentage increase in the number of <br> accounts services in a territory will increase <br> expected sales by $\$ 3,209$. |
| $\mathrm{X}_{3}$ | Percentage of US <br> personal income | -38.325 | A percentage increase in US personal <br> income in the territory will reduce expected <br> sales by $\$ 38,325$. |
| $\mathrm{X}_{4}$ | Percentage of non-white <br> population | -43.148 | A percentage increase in non-white <br> population in the territory will reduce <br> expected sales by $\$ 43,148$. |

(g) The actual dollar value of the regression residuals are calculated in (d) above and are depicted as "Deviations" in the table. However, analysis and comparison of residuals are difficult due to different sales levels in absolute terms. To facilitate analysis, the residuals are standardized in the scatterplot below and the residuals appear to be random.

Dependent Variable: Sales


Regression Standardized Predicted Value

Since there are only 11 territories as sample points, the studentized residuals are plotted below and the studentized residuals appear to be random as well.

Dependent Variable: Sales


Regression Standardized Predicted Value
(h) Based on the residual scatterplots in (g) above, any territories that exceed or fall below the sales norm by greater than one $S_{e}$ can be determined by observing for points above the +1 and below the -1 level on the regression residual scale. In both analysis, Territory 1 exceeds the sales norm by greater than one $S_{e}$ and Territory 9 lags the sales norm by greater than one $\mathrm{S}_{\mathrm{e}}$.
4. The men's track records for 55 countries are provided. Records for 8 events are included: $100 \mathrm{~m}, 200 \mathrm{~m}, 400 \mathrm{~m}, 800 \mathrm{~m}, 1,500 \mathrm{~m}, 5,000 \mathrm{~m}, 10,000 \mathrm{~m}$, and Marathon. The unit of measurement is minutes for all events except for $100 \mathrm{~m}, 200 \mathrm{~m}$ and 400 m which are in seconds. In order to obtain a consistent unit of measurement, all records in minutes are converted to seconds using the data transformation and compute feature of SPSS prior to the statistical analysis.

To ensure that it is appropriate to use factor analysis, the anti-image correlation matrix, Bartlett test of sphericity, and measure of sampling adequacy tests are conducted. The anti-image correlation matrix table below indicates that all variables have relatively low anti-image correlations with the other variables, indicating that the data is not unsuitable for factor analysis.

Anti-image Matrices

|  |  | V1 | V2 | V3 | V4S | V5S | V6S | V7S | V8S |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Anti-image Covariance | V1 | . 123 | -. 085 | -. 046 | -. 008 | . 013 | . 008 | -. 010 | . 016 |
|  | V2 | -. 085 | . 112 | -. 009 | -. 008 | -. 018 | -. 008 | . 008 | . 002 |
|  | V3 | -. 046 | -. 009 | . 155 | -. 038 | -. 008 | -. 001 | -. 002 | -. 002 |
|  | V4S | -. 008 | -. 008 | -. 038 | . 116 | -. 038 | -. 002 | . 002 | -. 012 |
|  | V5S | . 013 | -. 018 | -. 008 | -. 038 | . 073 | -. 009 | -. 015 | . 009 |
|  | V6S | . 008 | -. 008 | -. 001 | -. 002 | -. 009 | . 045 | -. 024 | -. 013 |
|  | V7S | -. 010 | . 008 | -. 002 | . 002 | -. 015 | -. 024 | . 033 | -. 026 |
|  | V8S | . 016 | . 002 | -. 002 | -. 012 | . 009 | -. 013 | -. 026 | . 095 |
| Anti-image Correlation | V1 | .836(a) | -. 729 | -. 333 | -. 066 | . 133 | . 108 | -. 159 | . 149 |
|  | V2 | -. 729 | .872(a) | -. 071 | -. 071 | -. 200 | -. 109 | . 126 | . 021 |
|  | V3 | -. 333 | -. 071 | .958(a) | -. 285 | -. 072 | -. 015 | -. 029 | -. 014 |
|  | V4S | -. 066 | -. 071 | -. 285 | .947(a) | -. 417 | -. 022 | . 031 | -. 109 |
|  | V5S | . 133 | -. 200 | -. 072 | -. 417 | .933(a) | -. 156 | -. 309 | . 109 |
|  | V6S | . 108 | -. 109 | -. 015 | -. 022 | -. 156 | .914(a) | -. 610 | -. 192 |
|  | V7S | -. 159 | . 126 | -. 029 | . 031 | -. 309 | -. 610 | .872(a) | -. 466 |
|  | V8S | . 149 | . 021 | -. 014 | -. 109 | . 109 | -. 192 | -. 466 | .934(a) |

The Bartlett's test of sphericity indicates that the correlations among the variables are significant. The MSA index of 0.909 also indicates high correlations among the variables.

| KMO and Bartlett's Test |  |
| :--- | ---: |
| Kaiser-Meyer-Olkin Measure of Sampling .909  <br> Adequacy.   <br>   719.113 <br> Bartlett's Test of Approx. Chi-Square 28 <br> Sphericity df .000 |  |

The first two principal components are required. Initially, the countries' location in terms of continents are coded and used as part of the factor analysis to determine if there is any relationship between geographical location of country and the records. However, it was found that the correlations are very low and we concluded that the geographical location of the country should not be used as part of the factor analysis. The correlation matrix of geographical location (V9C below) and the time records is provided below.

|  |  | V1 | V2 | V3 | V4S | V5S | V6S | V7S | V8S | V9C |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Correlation | V1 | 1.000 | . 923 | . 841 | . 756 | . 700 | . 619 | . 633 | . 520 | -. 133 |
|  | V2 | . 923 | 1.000 | . 851 | . 807 | . 775 | . 695 | . 697 | . 596 | -. 239 |
|  | V3 | . 841 | . 851 | 1.000 | . 870 | . 835 | . 779 | . 787 | . 705 | -. 120 |
|  | V4S | . 756 | . 807 | . 870 | 1.000 | . 918 | . 864 | . 869 | . 806 | -. 104 |
|  | V5S | . 700 | . 775 | . 835 | . 918 | 1.000 | . 928 | . 935 | . 866 | -. 194 |
|  | V6S | . 619 | . 695 | . 779 | . 864 | . 928 | 1.000 | . 975 | . 932 | -. 231 |
|  | V7S | . 633 | . 697 | . 787 | . 869 | . 935 | . 975 | 1.000 | . 943 | -. 230 |
|  | V8S | . 520 | . 596 | . 705 | . 806 | . 866 | . 932 | . 943 | 1.000 | -. 194 |
|  | V9C | -. 133 | -. 239 | -. 120 | -. 104 | -. 194 | -. 231 | -. 230 | -. 194 | 1.000 |
| Sig. (1-tailed) | V1 |  | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 | . 167 |
|  | V2 | . 000 |  | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 | . 039 |
|  | V3 | . 000 | . 000 |  | . 000 | . 000 | . 000 | . 000 | . 000 | . 192 |
|  | V4S | . 000 | . 000 | . 000 |  | . 000 | . 000 | . 000 | . 000 | . 224 |
|  | V5S | . 000 | . 000 | . 000 | . 000 |  | . 000 | . 000 | . 000 | . 078 |
|  | V6S | . 000 | . 000 | . 000 | . 000 | . 000 |  | . 000 | . 000 | . 045 |
|  | V7S | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 |  | . 000 | . 046 |
|  | V8S | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 |  | . 078 |
|  | V9C | . 167 | . 039 | . 192 | . 224 | . 078 | . 045 | . 046 | . 078 |  |

Principal component analysis was performed to provide a general outline of the relationships. Factor analysis using the principal components approach results in one factor. As a result, the extraction methodology was changed so that 2 factors are extracted from the analysis. All the available extraction methodologies were performed
and the resulting factors are very similar. Output of the factor analysis using the principal components method is provided below.

Descriptive Statistics

|  | Mean | Std. Deviation | Analysis N |
| :--- | ---: | ---: | ---: |
| V1 | 10.4711 | .35143 | 55 |
| V2 | 20.9404 | .64465 | 55 |
| V3 | 46.4387 | 1.45702 | 55 |
| V4S | 107.5964 | 3.82109 | 55 |
| V5S | 221.8909 | 9.35456 | 55 |
| V6S | 830.7491 | 48.06963 | 55 |
| V7S | 1739.3455 | 108.46390 | 55 |
| V8S | 8197.4400 | 553.62201 | 55 |

Correlation Matrix

|  |  | V1 | V2 | V3 | V4S | V5S | V6S | V7S | V8S |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Correlatio <br> n | V1 | 1.000 | . 923 | . 841 | . 756 | . 700 | . 619 | . 633 | . 520 |
|  | V2 | . 923 | 1.000 | . 851 | . 807 | . 775 | . 695 | . 697 | . 596 |
|  | V3 | . 841 | . 851 | 1.000 | . 870 | . 835 | . 779 | . 787 | . 705 |
|  | V4S | . 756 | . 807 | . 870 | 1.000 | . 918 | . 864 | . 869 | . 806 |
|  | V5S | . 700 | . 775 | . 835 | . 918 | 1.000 | . 928 | . 935 | . 866 |
|  | V6S | . 619 | . 695 | . 779 | . 864 | . 928 | 1.000 | . 975 | . 932 |
|  | V7S | . 633 | . 697 | . 787 | . 869 | . 935 | . 975 | 1.000 | . 943 |
|  | V8S | . 520 | . 596 | . 705 | . 806 | . 866 | . 932 | . 943 | 1.000 |
| Sig. (1tailed) | V1 |  | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 |
|  | V2 | . 000 |  | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 |
|  | V3 | . 000 | . 000 |  | . 000 | . 000 | . 000 | . 000 | . 000 |
|  | V4S | . 000 | . 000 | . 000 |  | . 000 | . 000 | . 000 | . 000 |
|  | V5S | . 000 | . 000 | . 000 | . 000 |  | . 000 | . 000 | . 000 |
|  | V6S | . 000 | . 000 | . 000 | . 000 | . 000 |  | . 000 | . 000 |
|  | V7S | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 |  | . 000 |
|  | V8S | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 | . 000 |  |

Communalities

|  | Initial | Extraction |
| :--- | ---: | ---: |
| V1 | 1.000 | .950 |
| V2 | 1.000 | .939 |
| V3 | 1.000 | .892 |
| V4S | 1.000 | .900 |
| V5S | 1.000 | .938 |
| V6S | 1.000 | .965 |
| V7S | 1.000 | .973 |
| V8S | 1.000 | .943 |

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 | 6.622 | 82.777 | 82.777 | 6.622 | 82.777 | 82.777 | 4.186 | 52.323 | 52.323 |
| 2 | . 878 | 10.970 | 93.747 | . 878 | 10.970 | 93.747 | 3.314 | 41.424 | 93.747 |
| 3 | . 159 | 1.992 | 95.739 |  |  |  |  |  |  |
| 4 | . 124 | 1.551 | 97.289 |  |  |  |  |  |  |
| 5 | . 080 | . 999 | 98.288 |  |  |  |  |  |  |
| 6 | . 068 | . 850 | 99.137 |  |  |  |  |  |  |
| 7 | . 046 | . 580 | 99.717 |  |  |  |  |  |  |
| 8 | . 023 | . 283 | 100.000 |  |  |  |  |  |  |

Extraction Method: Principal Component Analysis.

Scree Plot


Component Number

## Component Matrix(a)

|  | Component |  |
| :--- | :---: | :---: |
|  | 1 |  |
| V1 | .817 | .531 |
| V2 | .867 | .432 |
| V3 | .915 |  |
| V4S | .949 |  |
| V5S | .959 |  |
| V6S | .938 |  |
| V7S | .944 |  |
| V8S | .880 | -.411 |

Extraction Method: Principal Component Analysis.
a 2 components extracted.

Rotated Component Matrix(a)

|  | Component |  |
| :--- | :--- | ---: |
|  | 1 | 2 |
| V1 |  | .935 |
| V2 |  | .893 |
| V3 | .543 | .773 |
| V4S | .712 | .627 |
| V5S | .813 | .525 |
| V6S | .902 |  |
| V7S | .903 |  |
| V8S | .936 |  |

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. a Rotation converged in 3 iterations.

Component Transformation Matrix

| Component | 1 | 2 |
| :--- | ---: | ---: |
| 1 | .759 | .651 |
| 2 | -.651 | .759 |

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

Based on the results of the factor analysis using principal components method, the two factors that are derived can be classified as 'Sprints" and "Runs." 'Sprints" includes the events of $100 \mathrm{~m}, 200 \mathrm{~m}$ and 400 m (variables V1, V2 and V3 respectively). "Runs" includes the events of $800 \mathrm{~m}, 1,500 \mathrm{~m}, 5,000 \mathrm{~m}, 10,000 \mathrm{~m}$ and the Marathon (variables V4S, V5S, V6S, V7S and V8S respectively). The distance of the race is the major difference between these two factors identified.

The two factors "Sprints" and "Runs" can be interpreted as categories of one criterion by which the results can be classified into meaning representations. For example, the time required for Sprints are generally under one minute while Runs take over one minutes. There are also strategic differences in that runners run as fast as they can in Sprints but runners might utilize strategic positioning during the race in Runs.


[^0]:    a Multiple modes exist. The smallest value is shown

[^1]:    a Predictors: (Constant), X1 Percentage of US Footware
    b Dependent Variable: Y Sales

