# A Multivariate Statistical Analysis of Stock Trends

April Kerby
Alma College
Alma, MI

James Lawrence
Miami University
Oxford, OH

## Abstract

Is there a method to predict the stock market? What factors determine if a company's stock value will rise or fall in a given year? Using the multivariate statistical methods of principal component analysis and discriminant analysis, we aim to determine an accurate method for classifying a company's stock as a good or a poor investment choice. Additionally, we will explore the possibilities for reducing the dimensionality of a complex financial and economic dataset while maintaining the ability to account for a high percentage of the overall variation in the data.

## Introduction

The stock market is a financial game of winners and losers. Is your stock tried and true or here today – gone tomorrow? How can one pick out a golden nugget like Microsoft from the hundreds of dot-comers that went bust after an all-too-brief moment of glory? Indeed, it may seem like there is really no way to tell – a seemingly uncountable number of variables influence our markets and companies; how can one take all of them into account? Is there a simpler way of looking at the market madness?

This paper seeks to use statistical methods to survey and analyze financial and economic data to discover such a method of simplification. Using Principal Component Analysis, we will combine related factors into a smaller number of key components largely responsible for the variations observed. Then, using Discriminant Analysis, we will develop a model for separating companies into two categories based on their predicted stock performance: good and poor investment choices.

The data we use in this analysis comes from the Federal Reserve Bank of St Louis, Big Charts Historical Stock Quotes, as well as each company's annual reports. We use four company-specific and six macroeconomic variables that we feel might account for a particular company, at a specific time, being a good or a poor investment. The ten variables are as follows: net revenue, net income, price per earnings ratio of stock, diluted earnings per share, consumer spending, consumer investment, unemployment rate, inflation rate, federal funds rate, and the Dow Jones industrial average.

## Assessment of Normality

In order to run the analysis tests, we first check the normality of the data.
Normality testing involves regressing the variables, centered about their mean, against the z-values of each variable. A strong linear relationship implies normal data. The sample correlation coefficient, r, is calculated by:

$$r = \frac{\operatorname{cov}(x_{(i)}, q_{(i)})}{\sqrt{(\operatorname{var}(x_{(i)})\operatorname{var}(q_{(i)})}},$$

where x is the data value and q is the associated normal quartile. The null hypothesis, $H_0$: $\rho = 1$, implies normality and is rejected if r is less than the critical value.

The original variables must be tested individually for univariate normality. Once the variables are tested (using a Q-Q plot), those that are not normal may be thrown out or kept depending on the discretion of the researcher. The variables which do not display a normal distribution may be transformed in order to achieve normality. The log function or the square root function may be applied to help obtain normal variables. For our data, the critical value for r is 0.9771, at the $\alpha = .01$ level of significance. The variables called transformed revenue, transformed earnings per share, transformed consumer spending, inflation and unemployment rate are found to be normal. Transformed investments and the percent growth of the DJIAM are rejected with r-values of .969 and .967, respectively. Although five variables failed to pass the normality test, this was to be expected of financial/economic data, and the robust characteristics of the tests allows us to continue with the analysis.

## Theory of Principal Component Analysis

Principal Component Analysis (PCA) reduces the dimensionality of the data set by linearly combining the original correlated variables into new variables, some of which are ignored. These new variables are linearly independent of one another, whereas the original variables may have been dependent. Reducing the dimensionality allows fewer variables to be used to obtain a similar amount of information, strengthening the results of any subsequent statistical tests used. Additionally, since the principal components are composed of multiple variables, they allow for a more accurate sense of how the variables are interacting.

After the variables have been tested for normality, the eigenvalues of the variance-covariance matrix $\Sigma$ corresponding to the original variables are calculated. Each eigenvalue $\lambda_i$, with $i$=1,…p, corresponds to a particular eigenvector. This vector of coefficients is then used in transforming the linear combination of x-variables into principal components. The ideal goal is to maximize the var($\underline{l}_i^{'}\,\underline{x}$), where $\underline{l}_i^{'}$ is the vector of eigenvalues and $\underline{x}$ represents the original vector. Generally, only eigenvectors with corresponding eigenvalues greater than one are used since smaller eigenvalues contribute little to the total variance. To determine which principal components to keep, the eigenvalue scree plot may be used. There is typically an *elbow* in the shape of the scree plot. It is customary to keep those principal components prior to the elbow. The percent of variation explained by an individual principal component is equal to the ratio $\frac{\lambda_i}{\sum_{j=1}^{p}\lambda_j}$, and the total variance explained by the first k principal components is $\frac{\sum_{i=1}^{k}\lambda_i}{\sum_{j=1}^{p}\lambda_j}$.

After determining how many principal components to use, the selected components, denoted $y_i$ for the $i^{th}$ component, are calculated. Individually, each $y_i$ equals $\sum_{j=1}^{p} l_{ij} x_j$, where $l_{ij}$ corresponds to the $j^{th}$ element of the $i^{th}$ eigenvector, and as a whole, $\vec{y} = L\vec{x}$. Each $y_i$ is orthogonal to the other principal components, ensuring linear independence within the principal component variables.

**Principal Component Analysis Application**

Before applying the PCA test, we must first check to see that the variables are, in fact, dependent. Therefore, the test for independence must be executed. Using the correlation matrix R we test $H_0$: P = I against $H_1$: P $\neq$ I. The likelihood ratio test is used to test $H_0$ at $\alpha$=.01. The null hypothesis is rejected if the test statistic is greater than the critical value. In our case the critical value is $\chi^2_{.45,.01} = 69.9569$. The test statistic is calculated with the formula

$$-(n-1-\frac{2p+5}{2})\ln|R|.$$

Our test statistic is 779.2479 and the P-value is $1.34627\text{x}10^{-134}$ which strongly suggests that $H_0$ be rejected. Correspondingly, we reject the null hypothesis since the test statistic is greater than the critical value. This confirms that our variables are in fact dependent, and principal component analysis may be carried out.

Principal component analysis, run on the data using Minitab, provides the results in Table 1.1. The corresponding eigenvalues, as well as the variance accounted for by each of the ten principal components are given.

Table 1.1: Principal Components

| | PC 1 | PC 2 | PC 3 | PC 4 | PC 5 | PC 6 |
|---|---|---|---|---|---|---|
| Eigenvalue | 3.2473 | 1.928 | 1.5332 | 1.1832 | 0.8721 | 0.6659 |
| Proportion | 0.325 | 0.193 | 0.153 | 0.118 | 0.087 | 0.067 |
| Cumulative | 0.325 | 0.518 | 0.671 | 0.789 | 0.876 | 0.943 |
| | | | | | | |
| | PC 7 | PC 8 | PC 9 | PC 10 | | |
| Eigenvalue | 0.3216 | 0.2187 | 0.0218 | 0.0081 | | |
| Proportion | 0.032 | 0.022 | 0.002 | 0.001 | | |
| Cumulative | 0.975 | 0.997 | 0.999 | 1 | | |

In this case, four principal components have eigenvalues greater than one and are kept, accounting for 78.9% of the total variation. Next, to help confirm the retention of only four principal components, the eigenvalue scree plot may be useful. However, in this particular case, the scree plot does not clearly determine the number of components to be used, and we will base our model on the eigenvalues themselves. The principal components are computed using the given coefficients.

Table 1.2: Principal Component Coefficients

| Variables | PC 1 | PC 2 | PC 3 | PC 4 |
|---|---|---|---|---|
| Log Revenues | 0.082 | -0.043 | 0.377 | 0.601 |
| Log Diluted Earnings | -0.199 | 0.219 | 0.481 | 0.375 |
| Log Net Income | -0.155 | -0.410 | -0.178 | 0.552 |
| Unemployment Rate | 0.096 | 0.545 | -0.342 | 0.213 |
| Federal Funds Rate | -0.408 | -0.320 | 0.237 | -0.243 |
| % Growth DJIAM | -0.534 | 0.023 | -0.070 | -0.01 |
| Log Consumer Investments | 0.455 | -0.247 | 0.216 | -0.097 |
| Log Consumer Spending | 0.500 | -0.170 | 0.156 | 0.005 |
| Inflation | -0.109 | -0.178 | 0.348 | -0.225 |
| sqrt(Price per Earnings) | 0.018 | -0.512 | -0.475 | 0.170 |

It can be seen that the first component is affected largely by the macroeconomic variables and could be renamed Overall Economic Condition. The second principal component cannot be easily named because there is no discernable pattern in the variables that affect it. The third principal component is comprised mostly of diluted earnings and the price per earnings ratio, and could hence be called Strength of Stock. The net income and revenues dominate the fourth principal component, and this component could be renamed Profitability.

**Theory of Discriminant Analysis**

A tool is needed to classify a multivariate data vector into one of two populations. Examples of uses for this tool would be to classify students as likely to succeed or fail in college or to classify an organ transplant patient as likely to survive or not. Discriminant analysis provides a rule for classifying observations from a multivariate data set into two or more populations.

In the case of two populations defined by $\Pi_1 \equiv N_p(\underline{\mu}_1, \Sigma_1)$ and $\Pi_2 \equiv N_p(\underline{\mu}_2, \Sigma_2)$, we can derive a classification rule that can be used to classify an element $\underline{x}$ into one of the populations. Each $\underline{x}$ is assumed to be p-variate normal. When the population parameters are unknown, as is often the case, one must obtain training samples for estimation of the mean and covariance of each population, as well as derive the classification rule.

The estimates of $\underline{\mu}_1, \underline{\mu}_2$, $\Sigma_1$ and $\Sigma_2$, are $\overline{\underline{x}_1}$, $\overline{\underline{x}_2}$, $S_1$ and $S_2$, respectively. If the covariance matrices for the populations, $\Sigma_1$ and $\Sigma_2$, are equal, then the common covariance matrix, $\Sigma$, is replaced by the pooled estimate, $S_p = \dfrac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2}$. The estimates $\overline{\underline{x}_1}$, $\overline{\underline{x}_2}$, $S_1$, $S_2$, and $S_p$ are unbiased estimators.

Next comes the classification of $\underline{x}$ into $\Pi_1$ or $\Pi_2$. We use an optimal linear discriminant function, $\underline{a}`\underline{x}$, where $\hat{\underline{a}} = S_p^{-1}(\overline{\underline{x}}_1 - \overline{\underline{x}}_2)$, to assign $\underline{x}$ into a population based on the decision rule:

Classify $\underline{x}$ into $\Pi_1$ if $\hat{\underline{a}}'\underline{x} > \frac{1}{2}(\overline{\underline{x}}_1 - \overline{\underline{x}}_2)' S_p^{-1}(\overline{\underline{x}}_1 - \overline{\underline{x}}_2)$

otherwise, classify $\underline{x}$ into $\Pi_2$ if $\hat{\underline{a}}'\underline{x} \le \frac{1}{2}(\overline{\underline{x}}_1 - \overline{\underline{x}}_2)' S_p^{-1}(\overline{\underline{x}}_1 - \overline{\underline{x}}_2)$

After we have determined the decision rule, we calculate the apparent error rate (AER) based on the training sample used to derive the decision rule. The AER is the percentage of observations in the training sample that are misclassified by the decision rule. However, since the AER uses the specific data observations that were used to create the decision rule, there is a chance that the linear discriminant function and its corresponding decision rule will have a higher error rate in practice. The total probability of misclassification (TPM) denotes the probability that an individual observation will be misclassified using the derived function. This is related to the Mahalanobis distance between the two populations, $\Delta_p^2$, calculated by:

$$\Delta_p^2 = \underline{a}'\Sigma^{-1}\underline{a} = (\underline{\mu}_1 - \underline{\mu}_2)'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2)$$

Using this, we calculate:

$$\text{TPM} = 2\Phi(-\frac{1}{2}\Delta_p),$$

where $\Phi(x)$ is the cumulative distribution function of the standard normal distribution. For a data sample, TPM is well approximated by:

$$\hat{\alpha} = 2\Phi(-\frac{1}{2}\hat{\Delta}_p),$$

where $\hat{\Delta}_p^2 = (\overline{\underline{x}}_1 - \overline{\underline{x}}_2)' S_p^{-1}(\overline{\underline{x}}_1 - \overline{\underline{x}}_2)$.

Additionally, the Mahalanobis distance can be used as a classification rule equivalent to that of the linear discriminant function. If the Mahalanobis distance ($D_1^2$)

from $\underline{x}$ to $\Pi_1$ is less than the Mahalanobis distance ($D_2^2$) from $\underline{x}$ to $\Pi_2$, then we assign $\underline{x}$ to $\Pi_1$. Otherwise, we assign $\underline{x}$ to $\Pi_2$.

When using a linear discriminant function, it is assumed that $\Sigma_1 = \Sigma_2$. If this hypothesis is rejected, then we must use a quadratic discriminant function.

**Discriminant Analysis Application**

Before performing discriminant analysis, we first need a method of classifying a company as a good or poor investment, for a given year. While there is no definitive method for defining a market investment as "good" or "poor," here is a method that is simple and objective: if the value of a company's stock over a given year rose, it is classified as a good investment and otherwise it is classified as a poor investment. To obtain the stock prices at the end of each calendar year, we used the Big Charts database online [16]. Our training sample was based on a random selection of 15 companies, for all years from 1995-2002, where data was provided in their annual reports. This made a sample size of 88 distinct company-year observations. To create a test sample, we removed all eleven of the 2002 entries as well as seven randomly selected entries from the 1995-2001 data. The remaining 70 entries were used as the training sample.

First, it was necessary to test the assumption for using a linear discriminant function that $\Sigma_1 = \Sigma_2$. Using the likelihood ratio test, we tested $H_0$: $\Sigma_1 = \Sigma_2$ against $H_1$: $\Sigma_1 \neq \Sigma_2$. Our test statistic, $\chi^2_{obs} = \frac{m}{c}$, was tested against $\chi^2_{.45,.01} = 69.9569$. We obtained a value of 112.6807 for $\chi^2_{obs}$ with

$$m = \{\sum_{i=1}^{2}(n_i - 1)\}\{\ln\frac{|S_p|}{|S_i|}\}, \text{ and}$$

$$\frac{1}{c} = 1 - \frac{2p^2+3p-1}{6(p+1)}\left(\frac{1}{n_1-1} + \frac{1}{n_2-1} - \frac{1}{n_1+n_2-2}\right),$$

with a P-value of approximately 0. This is strong evidence that we should reject $H_0$. Therefore, we cannot use a linear discriminant model to classify the data; instead we must use a quadratic discriminant model. Initially, we used the principal components to attempt to create the discriminant model (see Table 1.3).

Table 1.3: Quadratic Model (Principal Components)

| | True Group Poor | True Group Good |
|---|---|---|
| Classified into Group | | |
| Poor | 16 | 4 |
| Good | 22 | 28 |
| | | |
| Total N | 38 | 32 |
| N Correct | 16 | 28 |
| Proportion | 0.421 | 0.875 |
| | | |
| N =70 | N correct = 44 | Proportion Correct = 0.629 |

However, this model produces consistently high error rates. Most notably, it correctly labels the poor group only 42.1% of the time and has an overall error rate of 37.1%. Using the ten variables resulting from the first set of transformations (see Appendices) produces better results, and therefore is kept for the final analysis. The following table summarizes the results of the quadratic model's classification of our training sample.

Table 1.4: Quadratic Model (Original Variables)

| | True Group Poor | True Group Good |
|---|---|---|
| Classified into Group | | |
| Poor | 36 | 15 |
| Good | 2 | 17 |
| | | |
| Total N | 38 | 32 |
| N Correct | 36 | 17 |
| Proportion | 0.947 | 0.531 |
| | | |
| N =70 | N correct = 53 | Proportion Correct = 0.757 |

The quadratic model has an apparent error rate (AER) of 24.3% percent. The theoretical TPM for this model was 34.8%. Considering the highly volatile nature of economic data, and the stock market in particular, this is regarded as a low error rate. However, as the purpose of the model was to predict how a stock would perform in the upcoming year, when the result is unknown, it is necessary to test our quadratic model on a test sample. For the test sample of 2002 data, plus seven random selections from 1995-2001, the discriminant model correctly classified 13 of the 18 stocks as a good or poor investment. The following table shows how the quadratic model classified the 18 stocks in our "unknown" test sample.

Table 1.5: Classification of Test Sample by Quadratic Model

| Observation | Stock | Predicted Group | Actual Group |
|---|---|---|---|
| 1 | Walt Disney 2002 | Good | Good |
| 2 | UPS 2002 | Good | Good |
| 3 | Tyson 2002 | Good | Poor |
| 4 | Motorola 2002 | Good | Good |
| 5 | Microsoft 2002 | Good | Poor |
| 6 | Kellogg 2002 | Good | Good |
| 7 | JCPenney 2002 | Good | Poor |
| 8 | Intel 2002 | Good | Good |
| 9 | Ford 2002 | Good | Good |
| 10 | Computer Associates 2002 | Good | Good |
| 11 | Boeing 2002 | Good | Good |
| 12 | Walt Disney 1995 | Good | Good |
| 13 | UPS 2000 | Poor | Poor |
| 14 | Motorola 2000 | Poor | Poor |
| 15 | Kellogg 2001 | Poor | Good |
| 16 | IBM 1999 | Poor | Poor |
| 17 | Ford 2001 | Good | Poor |
| 18 | Boeing 1998 | Good | Good |

The quadratic model placed 13 out of the 18 test stocks from outside the training sample into the correct population. The error rate of 27.7% means that this model correctly predicts whether the value of a particular stock will rise or fall over the next year nearly three out of four times – in times like these, enough to make a small fortune!

## Conclusion

We are satisfied with our last discriminant model because it predicts a high percentage of the outcomes for various stocks, and does not lose much accuracy when applied to a sample from outside the training sample. However, despite accounting for a high percentage of the overall variation, the four principal component model is unable to generate an accurate discriminant model for classification of a stock as a good or poor investment. Subsequent tests reveal that the PCA-based models do not achieve high accuracy until most of the components are employed in the discriminant model, which defeats the purpose of using the PCA data. Hence, we are unable to reduce the dimensionality of the data set. We believe that our discriminant model still has a great amount of room for improvement. Such improvement can be attained by adding additional variables, particularly those showing aspects of the company not related to profitability or earnings-share relationships. We might want to consider company size, distribution, or even the integration of online sales and marketing as supplementary factors. Additional global economic variables, such as GDP, trends in the Fed's purchase/sale of securities, and variables affecting world trade could provide additional clarity. As these variables are added, and perhaps with the usage of transformations not employed in this model, it may still be possible to reduce the dimensionality of the data set using PCA.

We believe that multivariate statistics can be used to analyze a company's likely performance in the stock market with respect to global economic conditions as well as its own financial performance the previous year.

**Acknowledgements**

We would like to thank Dr. Vasant Waikar for his tremendous help in mentoring our research. Without his guidance and teaching, none of this would have been possible. We would also like to thank Candace Porter, for her support and dedication. Additionally, we would like to acknowledge Dr. Marc Rubin and Dr. Larry Rankin for their expertise in the area of corporate finance. Finally, we would like to thank all of our professors, graduate assistants, and fellow SUMSRI students for their support, expertise, and camaraderie throughout the duration of the program.

**Bibliography**

[1] 3M Annual Reports (1995-2001). Retrieved June 27, 2003 from http://www.3m.com/US/

[2] Boeing Annual Reports (1995-2002). Retrieved June 27, 2003 from http://www.boeing.com/flash.html

[3] Computer Associates Annual Reports (1995-2002). Retrieved June 27,2003 from http://www.cai.com/

[4] Ford Motor Company Annual Reports (1998-2002). Retrieved June 28,2003 from http://www.ford.com/en/company/default.htm http://www.shareholder.com/ups/

[5] Hershey's Annual Reports (1997-2001). Retrieved June 28, 2003 from http://www.hersheyinvestorrelations.com/ireye/ir_site.zhtml?ticker=hsy&script=2100

[6] IBM Annual Reports (116-2001). Retrieved June 28, 2003 from http://www.ibm.com/us/

[7] Intel Annual Reports (1998-2002). Retrieved June 28,2003 from http://www.intel.com/

[8] JCPenney Annual Reports (1999-2002). Retrieved June 27, 2003 from http://www.jcpenney.net/

[9] Kelloggs Annual Reports (1998-2002). Retrieved June 28, 2003 from http://www.kelloggs.com/kelloggco/index.html

[10] Microsoft Annual Reports (1996-2002). Retrieved June 27, 2003 from http://www.microsoft.com/msft/ar.mspx

[11] Motorola Annual Reports (1998-2002). Retrieved June 27, 2003 from http://www.motorola.com/

[12] Tyson Foods, Inc. Annual Reports (1995-2002). Retrieved June 28, 2003 from http://www.tysonfoodsinc.com/

[13] UPS Annual Reports (2000-2002). Retrieved June 29, 2003 from http://www.shareholder.com/ups/

[14] Walmart Annual Reports (1996-1999). Retrieved June 25, 2003 from http://www.walmart.com/cservice/aw_index.gsp

[15] Walt Disney Annual Reports (1995-2002). Retrieved June 25, 2003 from http://disney.go.com/corporate/investors/shareholder/faq_investmentplan.html

**Bibliography (continued)**

[16] Big Charts Historical Quotes. (2003). Retrieved June 30, 20003 from http://bigcharts.marketwatch.com/historical/

[17] Federal Reserve Bank of St. Louis Economic Database. (2003). Retrieved June 26, 2003 from http://research.stlouisfed.org/fred/

[18] Johnson, Richard. Wichern, Dean. *Applied Multivariate Statistical Analysis: Fourth Edition*. Prentice Hall, Inc: New Jersey. 1998.

**Appendix 1.A** Data for 3M

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2001 | 10.21277337 | 0.553883027 | 5.746264271 | 9.477844 | 4.8 |
| 2000 | 10.21758921 | 0.553883027 | 5.801656477 | 9.33965 | 4 |
| 1999 | 10.19722541 | 0.63748973 | 4.74899941 | 9.459392 | 4.2 |
| 1998 | 10.17880435 | 0.459392488 | 4.969699857 | 9.29048 | 4.5 |
| 1997 | 10.17992503 | 0.704150517 | 4.027082232 | 9.536558 | 4.9 |
| 1996 | 10.15518416 | 0.558708571 | 4.788337582 | 9.394277 | 5.4 |
| 1995 | 10.13084818 | 0.36361198 | 5.360590521 | 9.336059 | 5.6 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Good |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Poor |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Good |
| 1998 | 3.762206 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Good |
| 1997 | 3.734312 | 3.725464863 | 1.571339 | 5.46 | 16.0749145 | Poor |
| 1996 | 3.719124 | 3.685810683 | 3.044041 | 5.3 | 19.51698701 | Poor |
| 1995 | 3.705487 | 3.647051254 | 2.727878 | 5.84 | 33.12260985 | Good |

**Appendix 1.B** Data for Boeing

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 10.73319727 | 0.457881897 | 3.390394301 | 9.502427 | 5.78 |
| 2001 | 10.76490806 | 0.532754379 | 3.372303963 | 9.551938 | 4.8 |
| 2000 | 10.71029511 | 0.387389826 | 5.200882649 | 9.476976 | 4 |
| 1999 | 10.76337558 | 0.396199347 | 4.079530645 | 9.521661 | 4.2 |
| 1998 | 10.7493807 | 0.06069784 | 5.32671691 | 9.145196 | 4.5 |
| 1997 | 10.66086548 | 0.744727495 | -16.48905361 | -8.53275 | 4.9 |
| 1996 | 10.3556622 | 0.267171728 | 5.365052077 | 9.394452 | 5.4 |
| 1995 | 10.29036856 | 1.397940009 | -31.30095845 | 8.556303 | 5.6 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Good |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Poor |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Good |
| 1998 | 3.762206 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Good |
| 1997 | 3.734312 | 3.725464863 | 1.571339 | 5.46 | 16.0749145 | Poor |
| 1996 | 3.719124 | 3.685810683 | 3.044041 | 5.3 | 19.51698701 | Poor |

| | | | | | | |
|---|---|---|---|---|---|---|
| 1995 | 3.705487 | 3.647051254 | 2.727878 | 5.84 | 33.12260985 | Good |

**Appendix 1.C** Data for Computer Associates

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 9.471878199 | -0.281033367 | -2.658582861 | -9.04218 | 5.78 |
| 2001 | 9.623042434 | -0.008600172 | -5.81495705 | -8.77159 | 4.8 |
| 2000 | 9.78554337 | 0.096910013 | 3.949683532 | 8.842609 | 4 |
| 1999 | 9.720407401 | 0.045322979 | 7.937821427 | 8.796574 | 4.2 |
| 1998 | 9.673849977 | 0.31386722 | 4.549085046 | 9.067815 | 4.5 |
| 1997 | 9.606381365 | -0.193820026 | 9.100137362 | 8.563481 | 4.9 |
| 1996 | 9.544688022 | 0.13667714 | -6.740797888 | -7.75092 | 5.4 |
| 1995 | 9.418798291 | 0.23299611 | 3.844948079 | 8.635387 | 5.6 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Good |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Good |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Poor |
| 1998 | 3.754631 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Good |
| 1997 | 3.734312 | 3.725464863 | 1.571339 | 5.46 | 16.0749145 | Poor |
| 1996 | 3.719124 | 3.685810683 | 3.044041 | 5.3 | 19.51698701 | Good |
| 1995 | 3.705487 | 3.647051254 | 2.727878 | 5.84 | 33.12260985 | Good |

**Appendix 1.D** Data for Ford

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 11.21330521 | 0.26760624 | -4.149966533 | -8.99123 | 5.78 |
| 2001 | 11.20822362 | -0.480006943 | -2.281512221 | 9.736635 | 4.8 |
| 2000 | 11.23061239 | 0.361727836 | 3.192382237 | 9.539954 | 4 |
| 1999 | 11.20590236 | 0.767897616 | 2.226891189 | 9.859559 | 4.2 |
| 1998 | 11.1563977 | 1.249442961 | 1.354006401 | 10.34382 | 4.5 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Good |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Poor |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Poor |
| 1998 | 3.762206 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Poor |

**Appendix 1.E** Data for Hershey

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2001 | 9.658701996 | 0.176091259 | 6.718134662 | 8.316298 | 4.8 |
| 2000 | 9.625412883 | 0.383815366 | 5.157839255 | 8.524452 | 4 |
| 1999 | 9.598891575 | 0.5132176 | 3.769021288 | 8.66305 | 4.2 |
| 1998 | 9.646953843 | 0.369215857 | 5.209294783 | 8.532612 | 4.5 |
| 1997 | 9.63369423 | 0.348304863 | 5.224983103 | 8.526664 | 4.9 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Good |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Good |
| 1998 | 3.754631 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Poor |
| 1997 | 3.734312 | 3.725464863 | 1.571339 | 5.46 | 16.0749145 | Good |

**Appendix 1.F** Data for IBM

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2001 | 10.93382123 | 0.638489257 | 5.273224493 | 9.887786 | 4.8 |
| 2000 | 10.94643261 | 0.64738297 | 4.375402169 | 9.907035 | 4 |
| 1999 | 10.94224623 | 0.614897216 | 5.117075925 | 9.886039 | 4.2 |
| 1998 | 10.9120466 | 0.81756537 | 5.297539706 | 9.801266 | 4.5 |
| 1997 | 10.87799297 | 0.778874472 | 4.17244746 | 9.784831 | 4.9 |
| 1996 | 10.88051062 | 0.699837726 | 5.499047277 | 9.73472 | 5.4 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Good |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Poor |
| 1998 | 3.754631 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Poor |
| 1997 | 3.734312 | 3.725464863 | 1.571339 | 5.46 | 16.0749145 | Good |
| 1996 | 3.719124 | 3.685810683 | 3.044041 | 5.3 | 19.51698701 | Poor |

**Appendix 1.G** Data for Intel

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 10.42755102 | -0.337242168 | 5.817888456 | 9.493737 | 5.78 |
| 2001 | 10.42388455 | -0.721246399 | 12.86570308 | 9.110926 | 4.8 |
| 2000 | 10.52796484 | 0.178976947 | 6.073773825 | 10.02263 | 4 |
| 1999 | 10.46818481 | 0.021189299 | 5.463733243 | 9.864155 | 4.2 |
| 1998 | 10.41950967 | -0.065501549 | 5.870699812 | 9.783046 | 4.5 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Good |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Poor |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Good |
| 1998 | 3.754631 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Good |

**Appendix 1.H** Data for JCPenney

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 10.50983401 | 0.136720567 | 4.09824602 | 8.607455 | 5.78 |
| 2001 | 10.50520426 | -0.585026652 | 10.17160452 | 7.991226 | 4.8 |
| 2000 | 10.50305489 | -0.44870632 | -1.967710883 | -9.75435 | 4 |
| 1999 | 10.51201697 | 0.064457989 | 4.146040903 | 8.526339 | 4.2 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Poor |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Good |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Poor |

**Appendix 1.I** Data for Kellogg

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 9.919287341 | 0.243038049 | 4.425252212 | 9.05854 | 5.78 |
| 2001 | 9.877831891 | 0.071882007 | 5.050591508 | 8.90531 | 4.8 |
| 2000 | 9.784403302 | 0.161368002 | 4.254814717 | 8.93837 | 4 |
| 1999 | 9.789369154 | -0.080921908 | 6.092658035 | 8.729732 | 4.2 |
| 1998 | 9.78604121 | 0.089905111 | 5.267633955 | 8.893484 | 4.5 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Good |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Good |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Good |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Poor |
| 1998 | 3.762206 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Poor |

**Appendix 1.J** Data for Microsoft

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 10.45278279 | 0.149219113 | 6.055300708 | 10.06119 | 5.78 |
| 2001 | 10.40305185 | 0.120573931 | 7.084447328 | 9.866051 | 4.8 |
| 2000 | 10.36089622 | 0.230448921 | 5.051499486 | 9.974097 | 4 |
| 1999 | 10.29550113 | 0.152288344 | 9.067431271 | 9.891259 | 4.2 |
| 1998 | 10.18361145 | -0.075720714 | 12.84940243 | 9.652246 | 4.5 |
| 1997 | 10.05530186 | -0.180456064 | 13.99404635 | 9.538322 | 4.9 |
| 1996 | 9.938069186 | -0.366531544 | 13.83877263 | 9.341435 | 5.4 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Poor |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Good |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Poor |
| 1998 | 3.754631 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Poor |
| 1997 | 3.734312 | 3.725464863 | 1.571339 | 5.46 | 16.0749145 | Good |
| 1996 | 3.719124 | 3.685810683 | 3.044041 | 5.3 | 19.51698701 | Good |

**Appendix 1.K** Data for Motorola

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 10.42616955 | -0.037426498 | -2.817051618 | -9.2584 | 5.78 |
| 2001 | 10.47527884 | -0.250420002 | -2.904858387 | -9.74123 | 4.8 |
| 2000 | 10.57495678 | -0.236572006 | 5.908789479 | 9.3485 | 4 |
| 1999 | 10.49039396 | -0.387216143 | 10.94108391 | 9.108227 | 4.2 |
| 1998 | 10.46683797 | 0.356547324 | -6.800735254 | -9.10721 | 4.5 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Good |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Poor |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Poor |
| 1998 | 3.762206 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Good |

**Appendix 1.L** Data for Tyson

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 10.36860296 | 0.033423755 | 3.223179934 | 8.773055 | 5.78 |
| 2001 | 10.02378728 | -0.397940009 | 5.373546315 | 7.944483 | 4.8 |
| 2000 | 9.861414919 | -0.173925197 | 4.362321715 | 8.178977 | 4 |
| 1999 | 9.882011962 | 0 | 4.031128874 | 8.361728 | 4.2 |
| 1998 | 9.870052582 | -0.958607315 | 13.89898623 | 7.39794 | 4.5 |
| 1997 | 9.803156556 | -0.070581074 | 4.910972109 | 8.267172 | 4.9 |
| 1996 | 9.809828961 | -0.22184875 | 7.555351304 | 7.934498 | 5.4 |
| 1995 | 9.741230411 | 0.178976947 | 4.159884104 | 8.340444 | 5.6 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Poor |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Poor |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Poor |
| 1998 | 3.754631 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Poor |
| 1997 | 3.734312 | 3.725464863 | 1.571339 | 5.46 | 16.0749145 | Good |
| 1996 | 3.719124 | 3.685810683 | 3.044041 | 5.3 | 19.51698701 | Poor |
| 1995 | 3.705487 | 3.647051254 | 2.727878 | 5.84 | 33.12260985 | Good |

**Appendix 1.M** Data for UPS

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 10.37883379 | 0.330413773 | 5.429238944 | 9.699751 | 5.78 |
| 2001 | 10.48174352 | 0.322219295 | 5.094347942 | 9.595165 | 4.8 |
| 2000 | 10.46979257 | 0.397940009 | 4.847679857 | 9.684307 | 4 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Good |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Good |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Poor |

**Appendix 1.N** Data for WalMart

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 1999 | 11.13872573 | -0.004364805 | 8.356029699 | 9.646404 | 4.2 |
| 1998 | 11.0717274 | -0.107905397 | 7.225311634 | 9.547282 | 4.5 |
| 1997 | 11.02060571 | 0.123851641 | 3.850593158 | 9.485153 | 4.9 |
| 1996 | 10.97140111 | 0.075546961 | 3.092414139 | 9.437751 | 5.4 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Poor |
| 1998 | 3.754631 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Good |
| 1997 | 3.734312 | 3.725464863 | 1.571339 | 5.46 | 16.0749145 | Good |
| 1996 | 3.719124 | 3.685810683 | 3.044041 | 5.3 | 19.51698701 | Good |

**Appendix 1.O** Data for Walt Disney

| Year | log Revenues | log Diluted EPS | sqrt PPE | log NI | Unemployment |
|---|---|---|---|---|---|
| 2002 | 10.40361804 | -0.22184875 | 5.213763836 | 9.340444 | 5.78 |
| 2001 | 10.40091772 | -0.958607315 | 13.01048528 | 9.108227 | 4.8 |
| 2000 | 10.40354945 | -0.244125144 | 8.191780219 | 9.420451 | 4 |
| 1999 | 10.36986496 | -0.207608311 | 6.475761258 | 9.364363 | 4.2 |
| 1998 | 10.36127442 | -0.050609993 | 5.339591366 | 9.499275 | 4.5 |
| 1997 | 10.35166105 | -0.022276395 | 9.212405823 | 9.529815 | 4.9 |
| 1996 | 10.27274641 | 0.292256071 | 5.6807049 | 9.314078 | 5.4 |
| 1995 | 10.08461202 | 0.414973348 | 4.697585304 | 9.325721 | 5.6 |

| Year | log CS | log CI | Inflation | Fed Funds Rate | % Growth DJIAM | Group |
|---|---|---|---|---|---|---|
| 2002 | 3.817962 | 3.799966723 | 2.597403 | 1.67 | -15.9109879 | Good |
| 2001 | 3.804289 | 3.813558306 | 1.142204 | 3.89 | -7.952341063 | Poor |
| 2000 | 3.794063 | 3.830432757 | 3.732227 | 6.24 | -1.404414789 | Poor |
| 1999 | 3.775574 | 3.696951061 | 2.738892 | 4.97 | 22.8480147 | Good |
| 1998 | 3.762206 | 3.772319506 | 1.670792 | 5.35 | 16.12470752 | Good |
| 1997 | 3.734312 | 3.725464863 | 1.571339 | 5.46 | 16.0749145 | Poor |
| 1996 | 3.719124 | 3.685810683 | 3.044041 | 5.3 | 19.51698701 | Good |
| 1995 | 3.705487 | 3.647051254 | 2.727878 | 5.84 | 33.12260985 | Good |

**Appendix 2.A** Linear Correlation Values for Normality Test

| Variables | Sample Correlation Coefficient |
|---|---|
| log Revenues | 0.986 |
| log Diluted Earnings | 0.991 |
| Log Net Income | 0.664 |
| Unemployment Rate | 0.982 |
| Federal Funds Rate | 0.921 |
| % Growth DJIAM | 0.967 |
| log Consumer Investments | 0.969 |
| log Consumer Spending | 0.982 |
| Inflation | 0.98 |
| sqrt(Price per Earnings) | 0.838 |

Note: Critical Value: r = .9771

Significance Level: $\alpha = .01$