A DISCRIMINANT FUNCTION ANALYSIS APPROACH TO STOCK SELECTION IN THE CAPITAL MARKET OF NIGERIA

Nwakuya Maureen.T. (PhD) &
Ofokansi Arch-a-Michael
Department of Mathematics and Statistics
University of Port Harcourt, **NIGERIA**

ABSTRACT

In this paper, an approach of selecting stocks in the capital market of Nigeria based on Discriminant Function Analysis is presented. Attempt was made to classify in two groups a sample of 20 companies similar in terms of business profile (manufacturing), that traded stocks on the Nigerian Stock Exchange during the period of January 2010 to December 2015 into high value and low value group (i.e. stocks that can generate more profit and less profit respectively). Initial grouping is made according to the Earnings per share ratio and splits the sample into '10' high value and '10' low value group. The analysis was performed on multivariate data of the companies and their six financial ratios. The findings of the study reveal that the Price per Earnings ratio and Closing price contributes the most to selecting stocks in the capital markets of Nigeria whereas Market capitalization contributes the least. Cross validation technique applied shows 85% prediction accuracy of generalizing the discriminant function model to predicting the group of new stocks of unknown membership. Thus, the Discriminant function derived in this paper can help potential investors in predicting group membership of newly listed company when selecting stocks that can generate more profit compared to others.

Keywords: Discriminate Functions, Financial Ratios, Cross Validation, Nigerian Stock Exchange.

INTRODUCTION

Investing as defined by Investopedia (2016) is "*the act of committing money or capital to an endeavor with the expectation of additional income or profit*". In the society today as a worker, there is a limit to how many hours a day you put in at work and you get paid based on the amount of time per day you put in at work. Making an investment can be achieved in several ways which include; putting money into securities (stocks), bonds, mutual funds, real estate, or most at times starting one's own business .etc. In Stock exchanges around the world, various types of securities are traded and stock belongs among them. Investing in stocks listed in the stock market involves the buying of shares issued by a company or combination of different companies across several sectors in this market. A useful and reliable way of indicating a company's performance and financial situation is by analyzing its financial ratios and compare with other firms with the same business profile. In the stock market, some common financial ratios include; Earnings per share, Return on Equity, Return on Assets, Payout ratio, Dividend yield, Dividend per share, Market Capitalization, Beta coefficients, Price per earnings ratio, Debt/Equity ratio.etc. When selecting stock that can give more profit which is seen as being of high value, situations arise where an investor will be interested in investigating the difference between all the companies alongside their financial ratios simultaneously. Thereby referring to the group of companies that can give more profit as "high valued stocks" and the ones that can give less profit as "low valued stocks".

An important Statistical technique that can achieve this is called Discriminant Function Analysis. Discriminant Analysis as its commonly called, is concerned with separating distinct sets of objects into groups and assigning new objects of unknown origin to one of the two (or more) distinct groups on the basis of measurement of some variables. Hafez et al. (2015). In this research, we will be concerned with the Two-group discriminant analysis. In general, high value stocks would have sound balance sheets that show good financial health. Jo Vu, (2011) observed that "*Financial ratios are often examined to supplement an investor's decision in equity markets*". The big problem for value investing will be in the decision making process, value minded investors are faced with the problem of selecting stocks based on their financial ratios that are of high-value so as to make more profit and diversify their portfolio. It is against this background that this study was conceived.

Aim and Objectives: The main aim of this study is to apply the method of discriminant analysis to classify and then predict stocks in the capital market of Nigeria on the basis of high value and low value.

LITERATURE REVIEW

Nature and Applications of Discriminant Analysis

Theoretically, discriminant analysis involves deriving a random variable (function) which is a linear combination of independent variables that will discriminate best between objects in the groups defined beforehand. It can also assign new objects to one group among a number of groups. Actually, two sets of techniques based on the purpose of analysis exist, i.e., Predictive discriminant analysis and Descriptive discriminant analysis.

Stevens (1996) described the distinction between Predictive and Descriptive discriminant analysis in the following way: *"in predictive discriminant analysis, the focus is on classifying subjects into one of several groups (or to predicate group membership) whereas in descriptive discriminant analysis, the focus is on revealing major differences among groups."* In their research, Oghojafor et al. (2012) developed a set of discriminatory functions which helped in predicting the willingness of subscribers to drop their current service provider. Okeke & Amobi (2014) employed discriminant analysis to classify retail bank customers on the basis of users and non users, and then they identified which variables contribute to the classification. Analysis of the data used for their study showed that the seven variables and individual factors variables analyzed significantly predicted group membership. In another study still under discriminant analysis, Hur-Yagba et al. (2015) examined the usefulness of multiple discriminant analysis model in analyzing the financial health of manufacturing companies in Nigeria. The study recommended that manufacturing companies in Nigeria should use the Altman discriminant model to help them detect signs of bankruptcy several years before it occurs.

Financial Ratios its predictive power

A significant number of studies like Jo Vu, 2011; Okicic et al. 2013; Armeanu et al. 2012; Stancu & Stancu, 2014, have shown that financial ratios have predictive power, which implies it can also be used for analyzing stocks.

Some Quantitative models in Stock Selection

Sorenson et al (2000) introduced an alternative approach to the traditional methods of stock screening based on a statistical technique known as Classification and Regression Tree (CART). Senol et al. (2012) proposed stock selection model based on fundamental and technical analysis by using artificial neural networks and support vector machines. Schadler and Eakins (2001) examined a stock selection model using Morningstar's style box.

Discriminant analysis in Stock Selection

Considering literature on researches embarked on using discriminant analysis to select stocks. Siquera et al. (2012) in their research analyzed the cross-sectional relation between fundamental and financial variables, besides the CAPM (Capital Asset Pricing Model) and the average stock return using discriminant analysis. They examined stocks traded on the Sao Paulo Stock exchange and found that the discriminarory predictive capacity obtained a significant level of success. In a different research, Okicic et al. (2013) proposed an approach of stock selection based on discriminant analysis by investigating some fundamental variables and average stock returns on the underdeveloped capital markets of Bosnia and Herzegovina.

DATA & METHODOLOGY

Data, Population, Sample

The study used secondary data. The secondary data was collected online through NSE contact center in which the data contained published annual reports from December 2010 – December 2015 on financial ratios for firms (companies) that issue stock in NSE. The population of this study comprises of companies similar in terms of business profile (manufacturing industry) that issued stocks traded on the Nigerian Stock Exchange between 2010 and 2015. The annual data on six financial ratios was obtained for 20 companies namely:

Glaxosmithkline Nigeria plc. **(GLAXOSMITH)**, *Nigerian Enamelware Company plc.* **(ENAMWELWA)**, *Flourmills plc.* **(FLOURMILL)**, *Beta Glass Company* **(BETAGLAS)**, *Cutix plc.* **(CUTIX)**, *Vitafoam Nigeria plc.* **(VITAFOAM)**, *PZ Cussons Nigeria plc* **(PZ)**, *Seven-Up Bottling company plc.* **(7UP),** *UAC OF Nigeria plc.* **(UACN)**, *Presco plc.* **(PRESCO)**, *Guiness Nigeria plc.* **(GUINESS)**, *Ashaka Cement plc.* **(ASHAKACEM),** *Dangote Sugar Refinery plc.* (DANGSUGA*R), S.C.O.A. Nigeria plc.* (SCOA), *Unilever Nigeria plc.* **(UNILEVER),** *Berger Paints plc.* **(BERGER),** *Capital Assets Limited* **(CAP),** *Aluminium Extrusion Industries plc.* **(ALEX),** *International Breweries plc.* **(INTBREW),** *Boc Gases Nigeria plc.* **(BOCGAS).**

The six financial ratios associated with the 20 samples include: *Price per Earnings ratio, Market Capitalization, Earnings per Share, Dividend Yield, Payout ratio, Closing price.*

Organization of Multivariate Data

Multivariate Data is a collection or measurement of p number of variables and n number objects/cases. That is p variables (j=1,2,…, p) and n objects/items (i=1,2,…,n) arranged in a rectangular array called **Y** of n rows and p columns. **Y** is called the data matrix where

$$\mathbf{Y} = \begin{array}{c} \\ \textit{object } 1 \\ \textit{object } 2 \\ \vdots \\ \textit{object } i \\ \vdots \\ \textit{object } n \end{array} \left[\begin{array}{ccccccc} \textit{variable } 1 & \textit{variable } 2 & \dots & \textit{variable } i & \dots & \textit{variable } j \\ y_{11} & y_{12} & \cdots & y_{1j} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2j} & \cdots & y_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ y_{i1} & y_{i2} & \cdots & y_{ij} & \cdots & y_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nj} & \cdots & y_{np} \end{array} \right]$$

Thus the annual report on the six financial ratios for each company is similar to a multivariate data matrix with *financial ratios* and *companies* representing *p variables* and *n objects* respectively which makes the data collected suitable for discriminant analysis.

Assumptions of Discriminant Analysis

- The variables Y_1 , Y_2 , … Y_P are independent of each other.
- Groups are mutually exclusive and the group sizes are not too different from each other.
- The number of independent variables is less than the sample size.
- The variance-covariance matrix of the independent variables are similar within each group of the dependent variable.
- The independent variables follow a normal distribution

Discriminant Analysis method

Johnson & Wichern (1992) observed that some situations may arise where one may be interested in separating two groups of objects or assigning a new object to one of two groups. It is convenient to label the groups G_1 and G_2 respectively.
The objects are separated or classified on the basis of measurements of p random variables $\mathbf{Y}^T = [Y_1, Y_2, \ldots, Y_p]$. The observed values of Y differ to some extent from one group to another.
Johnson & Wichern (1992) have suggested replacing the population parameters with their sample counterparts. Suppose we have n_1 observations of the multivariate random variable $\mathbf{Y}^T$ from G_1 and n_2 observations from G_2 with $n_1 + n_2 - 2 \geq p$
The respective data matrices are $\boldsymbol{Y}_{1\,(n_1 \times p)}$ and $\boldsymbol{Y}_{2\,(n_2 \times p)}$ where $(n_i \times p)$ indicates dimension of matrix i.e. n rows and p column, while $\boldsymbol{Y}_{1\,(n_1 \times p)}$ and $\boldsymbol{Y}_{2\,(n_2 \times p)}$ repesents group 2 repectively. The sample covariance matrices S_1 and S_2 can be combined (pooled) to derive a single unbiased estimate of $\sum$. In particular, the weighted average of the sample covariance matrices is computed as;

$$\mathbf{S}_{\mathbf{pl}} = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2}$$

$\mathbf{S}_{\mathbf{pl}}$ is an unbiased estimate of $\sum$ if the data matrices $\mathbf{Y_1}$ and $\mathbf{Y_2}$ contain random samples from the populations $\mathbf{G_1}$ and $\mathbf{G_2}$ respectively.Thus the linear discriminant function which maximally seperates two populations is given as

$$D = \left[\overline{Y}_1 - \overline{Y}_2\right] S_{pl}^{-1} Y$$

where Y is the random variable matrix(independent variables)
S_{pl}^{-1} is the inverse matrix of S_{pl}

In the next section we shall cover the descriptive and predictive aspect of discriminant analysis.

Descriptive Aspect of Discriminant Analysis

The importance of discriminant analysis in describing group differences can be accessed by considering its eigenvalue. Citing from Rencher (2002),

Let **H** represent the matrix in place of $\left[\overline{Y}_1 - \overline{Y}_2\right]\left[\overline{Y}_1 - \overline{Y}_2\right]^T$
E represent $\mathbf{S}_{\mathbf{pl}}$
and let **a** = $\left[\overline{Y}_1 - \overline{Y}_2\right] S_{pl}^{-1}$

ISSN 2056-6018

The solutions of $(\mathbf{E}^{-1}\mathbf{H} - \lambda\mathbf{I})\mathbf{a} = 0$ are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_s$ of $\mathbf{E}^{-1}\mathbf{H}$, so we rank them as $\lambda_1 > \lambda_2 > \cdots > \lambda_s$. The largest eigenvalue λ_1 after ranking is the maximum value of $\lambda = \boldsymbol{a}^T\boldsymbol{H}\boldsymbol{a}/\boldsymbol{a}^T\boldsymbol{E}\boldsymbol{a}$ that maximally seperates the means of each group and therefore describes group differences. An eigenvalue λ_1 that is greater than 1 indicates a good model.

We may wish to know how well the variables separate the groups. This can be gotten from the canonical correlation $\boldsymbol{r}$ which serves as a means of measuring the association between the groups in the dependent variable and the discriminant function.

Rencher (2002), derived the canonical correlation as

$$\boldsymbol{r} = \sqrt{\frac{\lambda_1}{1+\lambda_1}}$$

A high level of $\boldsymbol{r}$ within the range of values $(0 < \boldsymbol{r} < 1)$indicates a high level of association between the groups in the dependent variable and the discriminant function. Also the squared canonical correlation $\boldsymbol{r}^2 = \frac{\lambda_1}{1+\lambda_1}$ in discriminant analysis is analogous to the $\boldsymbol{R}^2$ coefficient of determination in regression analysis.

In order to test the significance of the discriminant function, we will apply Wilk's Λ test for *two group case*. As seen in Rencher (2002) p282,

$\Lambda = \frac{1}{1+\lambda_1}$, Λ ranges from the values 0 to 1 and the values of Λ closer to 0 indicates the significance of the discriminant function.

In the presence of other independent variables, we are interested in assessing the contribution from each of the variables. We standardize the discriminant function coefficients to offset the different scales among the independent variables.

Rencher (2002) p278, has shown that the relative contribution of each independent variables to separation of two groups can be suitably assessed by comparing the coefficients $\mathbf{a}_r$, r = 1,2, …, p in the discriminant function;

The standardized coefficient is of the vector form;

$$\boldsymbol{a}^* = (\text{diag}\mathbf{S}_{pl})^{1/2}\,\mathbf{a} \qquad r = 1,2,\ldots,p$$

where "**diag**" denotes "diagonal elements of"

The absolute value of these standardized coefficients is used to rank these variables in order of their contribution to separating groups.

We shall now proceed to consider the Predictive aspect of Discriminant analysis

Predictive aspect of Discriminant Analysis.

Suppose we have a new object to be classified into one of the two groups, the predictive aspect of discriminant analysis enables the allocation of new objects to groups. Consider an object whose group membership is unknown, we can assign it to a group on the basis of the p measured variables, Y_0, associated with this object. Applying Fisher (1936) classification procedure when there are two populations , a simple classification can be based on the discriminant function

$$D = a^TY = [\overline{\boldsymbol{Y}_1} - \overline{\boldsymbol{Y}_2}]^{\boldsymbol{T}}\mathbf{S}_{pl}^{-1}Y_0$$

Where Y_0 is the vector of measurement on a new object we wish to classify into one of the two groups (populations).

To determine whether Y_0 is closer to $\overline{\boldsymbol{Y}_1}$ or $\overline{\boldsymbol{Y}_2}$, we check if its discriminant score closer to the centroid $\overline{D_1}$ or $\overline{D_2}$.

We evaluate each object's disciminant score $D_{1i} = a^T Y_{1i}$ for each observation Y_{1i} from the first group and then $\overline{D_1} = \sum_{i=1}^{n_1} \frac{D1i}{n_1} = a^T \overline{Y_1}$ is centroid for first group. Similaarly we obtain $\overline{D_2}$ as centroid for second group.

Now suppose $\bar{Y}$ is the overall mean vector matrix for each p random variables (combining group 1 and 2). We can calculate its discriminant score $\bar{D} = a^T \bar{Y}$ as the threshold value.

Fisher (1936) linear classification procedure assigns;

- Y_0 to G_1 if $D_0 = a^T Y_0$ is closer to $\overline{D_1}$ than it is to $\overline{D_2}$ and
- Y_0 to G_2 if D_0 is closer to $\overline{D_2}$ than it is to $\overline{D_1}$

Cross Validation

Cross validation which is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set. For the purpose of this study, we shall apply the hold-out-one sample method which will be performed on times.

This method involves leaving out one object in the data and then perform discriminant analysis on the remaining n-1 objects to predict the membership of that sample that was left out. This process will be repeated n times leaving out a different sample each time.

At the end of the iteration, the prediction will be compared with the original classification to access the level of accuracy of the discriminate function of predicting new objects.

DATA ANALYSIS

This section presents the Discriminant Analysis results using IBM SPSS 21 and Minitab 17 software on our sample of 20 companies that are similar in terms of business profile (manufacturing) traded on the Nigerian Stock Exchange from 2010 – 2015.

Recall from the previous section that we will be analyzing the multivariate data of average annual financial ratios of listed companies from 2010 – 2015. For a sustainable characterization, we shall transform the Price per Earnings ratio, Market Capitalization and Closing Price by taking their natural logarithms respectively for each company as shown in Table1. Also since we are using Earnings per Share (EPS) ratio as the categorical dependent variable, the first ten companies with larger values of EPS ratio will be categorized under Group "1" (high value stocks) while the remaining ten companies will be categorized under Group "0" (low value stocks) as shown in Table 1 below.

For the purpose of this study, let

logMKTCAP represents (logarithmic)Market Capitalization

YDVD represents Dividend Yield

logCLPR represents (logarithmic) Closing Price

PAYR represents Payout ratio

logPE represents (logarithmic) Price per Earnings ratio

TABLE 1: ANNUAL 2010 to 2015 TRANSFORMED AVERAGE DATA OF FINANCIAL RATIOS FOR SELECTED COMPANIES BASED ON STOCK PRICES

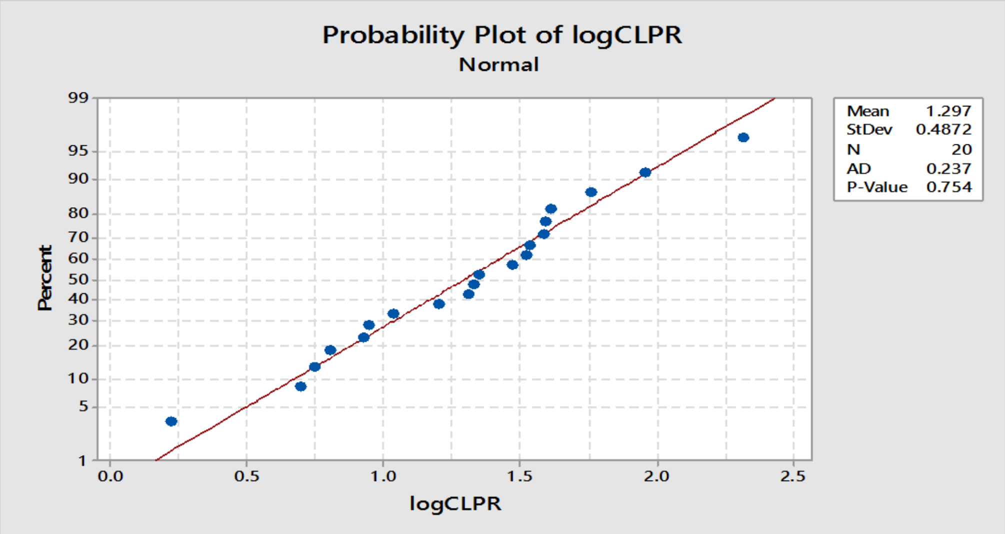
s/n	*firm*	logCLPR	*LogPE*	*YDVD*	*PAYR*	*logMKTCAP*	GROUP	*EPS*
1	GUINESS	2.3153	1.35	0.031	0.69	11.4879	1	9.17
2	ENAMELWA	1.5371	1.18	0.013	0.18	9.3390	1	7.61
3	UACN	1.5893	0.98	0.045	0.37	10.8357	1	4.84
4	FLOURMILL	1.7615	1.13	0.044	0.48	11.1080	1	4.76
5	PRESCO	1.3308	0.73	0.034	0.20	10.3308	1	4.67
6	7UP	1.9593	1.35	0.033	0.59	10.7658	1	3.82
7	BETAGLAS	1.3504	0.87	0.023	0.14	10.0494	1	3.22
8	GLAXOSMITH	1.6133	1.28	0.026	0.47	10.6089	1	2.30
9	CAP	1.5231	1.28	0.053	0.79	10.3004	1	1.80
10	UNILEVER	1.5934	1.47	0.021	0.61	11.1712	1	1.32
11	PZ	1.4710	1.40	0.019	0.46	11.0452	0	1.26
12	ASHAKACEM	1.3140	1.36	0.020	0.40	10.6642	0	1.15
13	BERGER	0.9446	0.92	0.078	0.64	9.3501	0	1.13
14	DANGSUGAR	0.9275	0.99	0.079	0.73	11.3224	0	0.84
15	BOCGAS	0.8043	1.00	0.036	0.40	9.4093	0	0.73
16	VITAFOAM	0.6949	0.84	0.060	0.41	9.6237	0	0.71
17	INTBREW	1.2058	1.55	0.011	0.33	10.7004	0	0.47
18	SCOA	0.7452	1.45	0.017	0.47	9.5580	0	0.35
19	ALEX	1.0338	1.72	0.006	0.33	9.3762	0	0.24
20	CUTIX	0.2231	1.01	0.074	0.74	9.3944	0	0.17

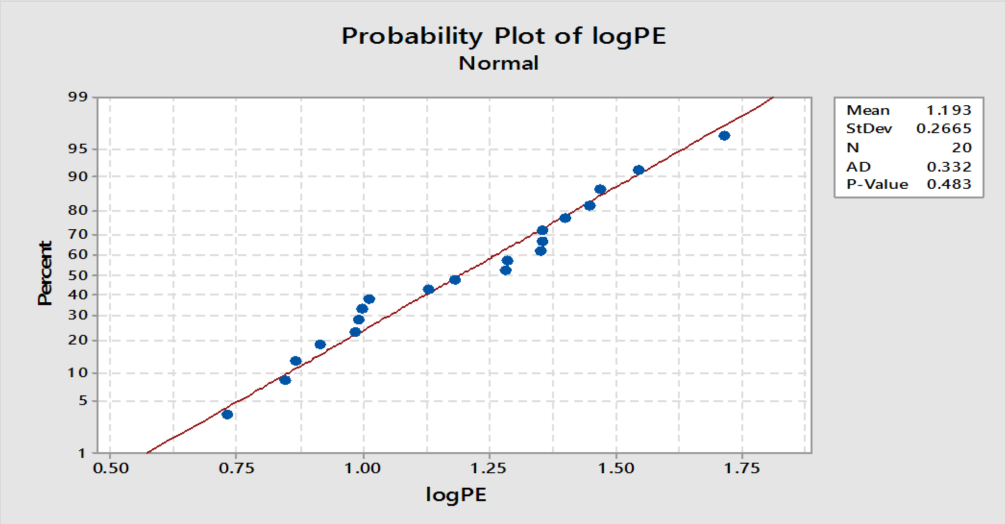
We first perform the test of equality of population covariance matrix using Box's M test, which tests the null hypothesis of homogenous covariance matrices. And if they are homogenous, the discriminante analysis provides a meaningful result.

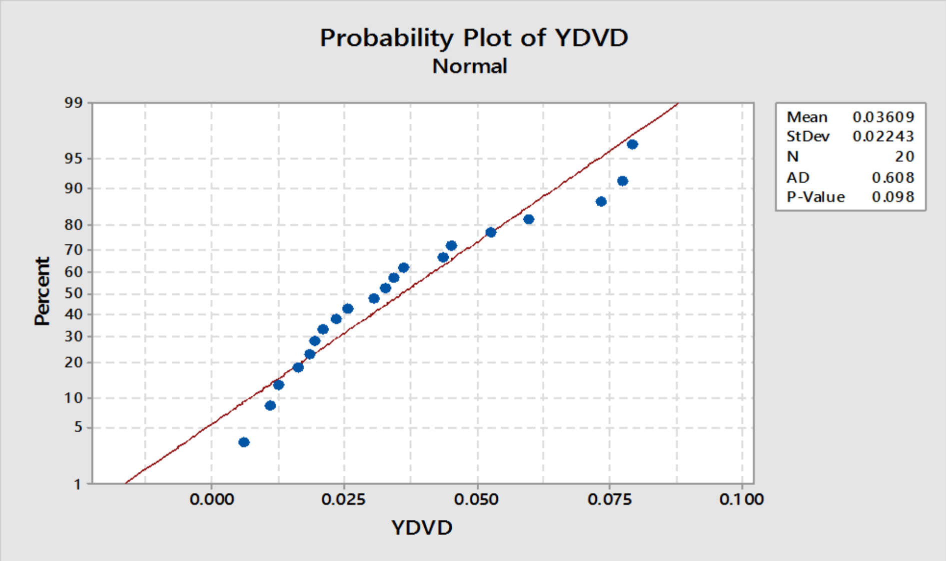
TABLE 2: Test of Covariance matrix

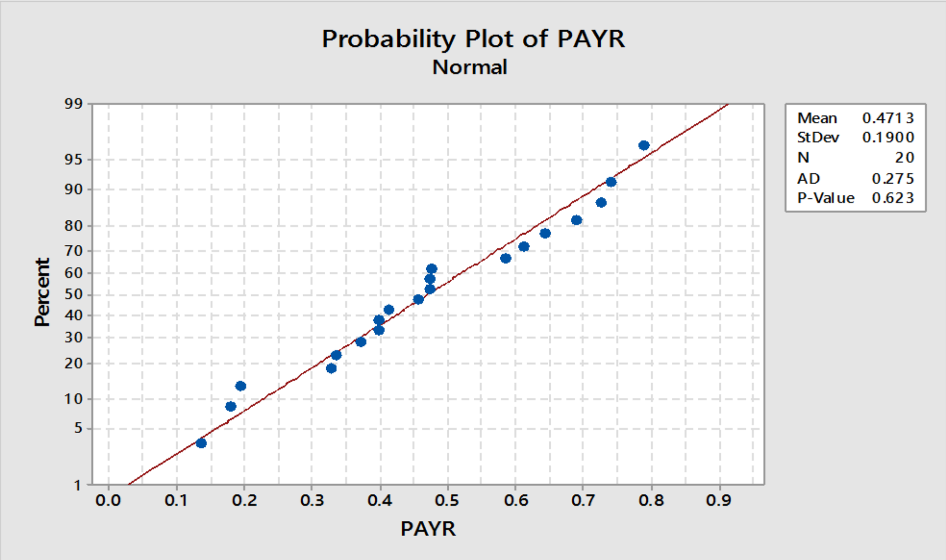
Box's M		20.731
F	Approx.	.957
	df1	15
	df2	1304.526
	Sig.	.500

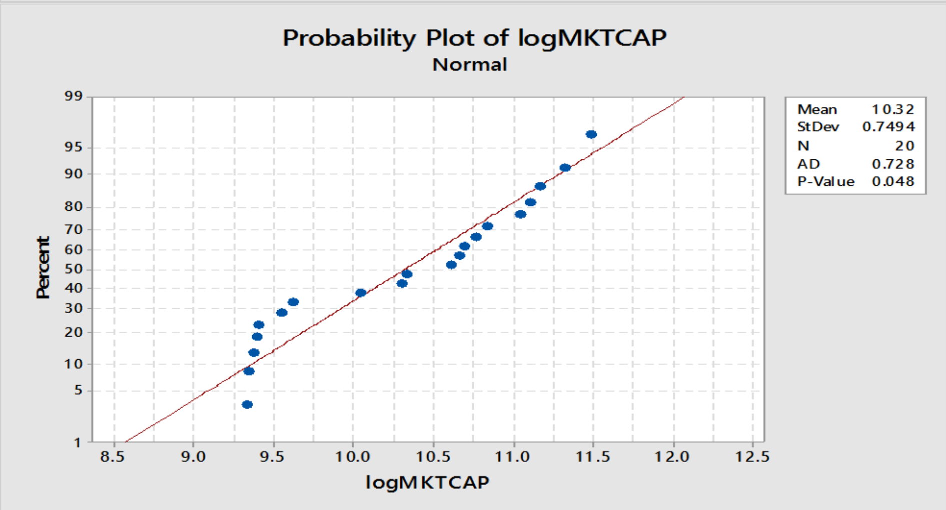
According to Table 3, Box's M test is not significant, since the p-value is greater than the level of significance 5% (ie. p-value = 0.50 > 0.05), therefore we accept the homogenous matrices hypothesis.


Probability Plot of logCLPR
Normal
Mean 1.297
StDev 0.4872
N 20
AD 0.237
P-Value 0.754
Percent
logCLPR
99
95
90
80
70
60
50
40
30
20
10
5
1
0.0
0.5
1.0
1.5
2.0
2.5


Probability Plot of logPE
Normal
Mean 1.193
StDev 0.2665
N 20
AD 0.332
P-Value 0.483
Percent
logPE
99
95
90
80
70
60
50
40
30
20
10
5
1
0.50
0.75
1.00
1.25
1.50
1.75







Testing the independent variables (financial ratios) for normality at α = 0.05 using Minitab 17 yields the above graphs and their respective p-values indicate they are normally distributed. Using SPSS 21 software yields Table 3 which indicates the eigenvalue, which measures the difference between the groups in the discriminant function. Observe that the eigenvalue is 2.870 (>1) indicating that the function is a good model.

TABLE 3: Eigen values and Canonical Correlation

Function	Eigen value	Canonical Correlation
1	2.870	.861

Also appearing in Table 3 above is the canonical correlation, which determines how much in percentage the function explains the discrimination between groups. We can achieve this by increasing the canonical correlation to its squared value. Thus $r^2 = 0.861^2 = 0.741$, i.e. the function explained 74.1% of the discrimination between groups

TABLE 4: Wilks' Lambda test

Test of Function(s)	Wilks' Lambda	Chi-square	Df	Sig.
1	.258	20.977	5	.001

Table 4 above shows that the discriminant function is statistically significant since p-value = 0.001 < 0.05, with Wilk's lambda being 0.258 which is closer to 0 than it is to 1. This indicates that the two groups; "High value" and "Low value" stocks seem to differentiate quite well.

Observe from Table 5 below;

Table 5: Standardized Canonical Discriminant Function Coefficients					
<u>*function:*</u>	***logCLPR***	***logPE***	***YDVD***	***PAYR***	***logMKTCAP***
	-1.367	1.609	1.194	-0.888	0.515

The absolute value of the standardized function coefficients shows that logPE, logCLPR, are the most important variables in selecting stocks with 1.609 and 1.367. The next important predictors are YDVD and PAYR. Notice also that logMKTCAP is the least important since it is the lowest with 0.515. Hence Price per earnings ratio and Payout ratio strongly predicts allocation of stocks to the high value or low value group in NSE, while Market capitalization has least effect in allocation.

Table 6: Canonical Discriminant Function Coefficients						
function:	***LogCLPR***	***LogPE***	***YDVD***	***PAYR***	***LogMKTCAP***	***(Constant)***
	-4.195	5.89	52.658	-4.584	0.724	-8.798

From Table 6 above which shows the Canonical Discriminant Function coefficients, the Discriminant function can be arranged as follows:

$$D = -8.798 + 0.724\boldsymbol{logMKTCAP} - 4.584\boldsymbol{PAYR} + 52.658\boldsymbol{YDVD} ++ 5.89\boldsymbol{logPE} - 4.195\boldsymbol{logCLPR}$$

We can determine the threshold score value using the canonical discriminant function, this threshold value will be used for the new classification of companies in the "high value" and "low value" groups. The threshold score is estimated by taking the average values of the financial ratios across companies and computing them in the function, i.e.

$$D = -8.798 + 0.724\boldsymbol{logMKTCAP} - 4.584\boldsymbol{PAYR} + 52.658\boldsymbol{YDVD} + + 5.89\boldsymbol{logPE} - 4.195\boldsymbol{logCLPR}$$

$$\bar{D} = -8.798 + 0.724\mathbf{(10.3376)} - 4.584\mathbf{(0.47)} + 52.658\mathbf{(0.036)} + 5.89\mathbf{(1.19)} - 4.195\mathbf{(1.2969)}$$

$$\therefore \bar{D} = -0.0038$$

ISSN 2056-6018

The threshold score is – 0.0038 and it will be used to separate companies between the two groups. After finding the discriminant function and the threshold score, one should perform centroid calculation of each group. The centroid which is the arithmetic mean of the discriminant scores of each group is shown below.

TABLE 7: Functions at Group Centroids

GROUP	Function
	1
0	1.607
1	-1.607

Thus from table 7, a stock is classified as belonging to the high value group if the calculation result of its discriminant score is lower than the threshold – 0.0038 (ie < – 0.0038). On the other hand, a stock is classified as belonging to the group of low value if the calculation result of its discriminant score is greater than the threshold – 0.0038 (ie. > – 0.0038).

Notice from Table 8 that the initial ordering of companies in Table 1 based on EPS variable is quite different from the new order in which the Discriminant score for each company was compared with the threshold value.

Table 8: Reclassification of 20 Companies

New order	***Firm***	***Old Order***	***Previous state***	***Current state***	***Discr. Score***
1	GUINESS	1	1	1	-4.61
2	7UP	6	1	1	-2.237
3	PRESCO	5	1	1	-1.727
4	ENAMWELA	2	1	1	-1.675
5	BETAGLAS	7	1	1	-1.493
6	FLOURMILL	4	1	1	-1.372
7	GLAXOSMITH	8	1	1	-1.304
8	UACN	3	1	1	-1.173
9	CAP	9	1	1	-0.98
10	UNILEVER	10	1	1	-0.426
11	PZ	11	0	0	0.166
12	BOCGAS	15	0	0	0.593
13	BERGER	13	0	0	0.603
14	ASHAKEM	12	0	0	0.641
15	VITAFOAM	16	0	0	1.484
16	INTBREW	17	0	0	2.087
17	DANSUGAR	14	0	0	2.115
18	SCOA	18	0	0	2.278
19	ALEX	19	0	0	2.588
20	CUTIX	20	0	0	3.523

Cross Validation classification

In this analysis we shall use cross validation from SPSS 21 to test how accurately our predictive model (discriminant function) will perform in practice. Table 9 presents the cross validated results.

Table 9: Cross Validated Classification Results

		GROUP	Predicted Group Membership		Total
			0	1	
Cross-validated	Count	0	8	2	10
		1	1	9	10
	%	0	80.0	20.0	100.0
		1	10.0	90.0	100.0

It can be observed that with the Cross Validated for the total sample of 20 cases, 17 (85%) overall are correctly classified. Note that this percentage is similar to the coefficient of determination, R^2 , in the regression model. Of the High value group 90% are correctly identified. Also, 80% of the Low value group are correctly classified.

This cross validation prediction of group membership provides a summary of how well the analysis would be at classifying new stocks that have not been included in the original sample of companies. So far, we can deduce that the discriminant analysis validates the intial grouping of stocks according to the categorical dependent variable EPS before we started the analysis.

CONCLUSION AND RECOMMENDATION

So far, we have investigated the possibility of applying Discrminant analysis to Stock selection and analysis on the capital markets of Nigeria. The results of the analysis shows that the six financial ratios applied together with the discriminant function model are significant in selecting stocks. Thus the main aim of this paper was to develop a disciminant function model that can be used to predict the group membership of a new manufacturing company that issue stock, when a potential investor is selecting which stock to buy. Note that the High value group of stocks can generate more profit. Cross validation prediction accuracy of 85% clearly indicates that the model can be reliably generalized to companies of unknown group membership. The following recommendation are made;

- Business managers should appreciate the importance of frequently subjecting their financial statements to ratio analysis to determine the financial health and profitability of their company.
- We recommend the use of financial ratios and Discriminant function analysis in selecting stocks.
- For further study in the capital markets of Nigeria, the method of Discriminant analysis can be applied to selecting stocks in other sectors such as; Oil, Health, Utilities, Banking, etc. Also more financial ratios can be added for the purpose of the study.

REFERENCES

Akpan, J. T (2013). The Extent of Relationship between Stock Market Capitalization and Performance on the Nigerian Economy. *Research Journal of Finance and Accounting.* ISSN 2222-1697(paper) ISSN 2222-2487(online) Vol.4, Issue19, 2013.

Erimafa, J.T., Idseri, A. and Edokpa I.W. (2009). Application of Discriminant Analysis to predict the class of degree for graduating students in a University System. *International Journal of Physical Sciences*. Vol.4 , Issue1, January 2009. pp016-017 ISSN 1992-1950 © 2009 Academic Journals, www.academicjournals.org/IJPS

Fisher, R.A. (1936). "The Use of Multiple Measurements in Taxonomic Problems", Annuals of Eugenics,7, pp 179-188.

Hafez, E.J., Abdel-Nabl S.M., Adel-Fatah, E.M. and Zeldan, A.S. (2015). Discriminant Analysis in View of Statistical and Operations Research Techniques. *Journal of Multidisciplinary Engineering Science and Technology. (JMEST).* ISSN: 3159-0040. Vol.2. Issue 11, November 2015.

Hur-Yagba, A.A., Okeji, I.F., Ayuba, B. (2015). Analyzing Financial Health of Manufacturing Companies in Nigeria Using Multiple Discriminant Analysis. *International Journal of Managerial Studies and Research (IJMSR).* Vol.3, Issue7, July 2015. pp 72-81. ISSN 2349-0330 (print) & ISSN 2349-0349(online) www.arcjournals.org.

Investopedia web site accessed on August 2016 via http://www.investopedia.com/term/v/stock_investing

Jaba E., Robu I.-B. (2009), Utilizarea analizei discriminant pentru obtinerea probelor de audit (I şi II), *Revista de audit financiar nr.11 şi 12 (2009),* https://www.researchgate.net/publication/235996415_Utilizarea_analizei_discriminant_pentru_obtinerea_probelor_de_audit_(I)

Johnson, A. R. & Wichern, W. D, (1992). Applied Multivariate Statistical Analysis. Prentice Hall, Englewood Cliff. New Jersey 07632. ISBN 0-13-041773-4.

Jo Vu, (2011). Discriminant Analysis and Identification of High Value Stocks. School of Accounting and Finance, Victoria University, Australia.pp1-4.

Lee, W., Tzeng, G., Guan, J., Chen, K., & Huang, J., (2009). Combined MCMD techniques for exploring stock selection based on Gordon model. *Expert Systems with Applications.* Vol.36, Issue 3. ISSN: 6421-6430.

Oghojafor, B., Mesike, G., Bakarae, R., Omoera, C., & Adeleke, I. (2012). Discriminant Analysis of Factors Affecting Telecoms Customer Churn. *International Journal of Business Administration*. Vol.3, Issue 2, March 2012. www.sciedu.ca/ijba, http://dx.doi.org/10.5430/ijba.v3n2p59

Okeke, T.C., & Amobi, D.C. (2014). A Discriminant Analysis of Electronic Banking in Nigeria. *Journal of Emerging Trends in Economics & Management Science. (JETEMS).* Vol. 5, Issue 2. pp 194-200. © Scholarlink Research Institute Journals, 2014. (ISSN: 2141- 7024), jetems.scholarlinkresearch.org

Okicic, J., Remetic, H.S, Buyukdemir, B., (2013). Stock Selection Based on Discriminant Analysis: Case of capital Markets of Bosnia and Herzegovina. Journal of Economics and Social Studies. Vol. 4, Issue 2. Fall, 2014. http://dx.doi.org/10.14706/JECOSS11421

Pandey, I.M. (2006). "Financial Management" -9th ed. Vilka Publisher, India.

Rencher, A.C. (2002). Methods of Multivariate Analysis -2nd ed. A Wiley-Interscience Publication. ISBN: 0-471-41889-7

Schadler, F.P. & Eakins, S.G. (2001). A Stock Selection Model using Morning Star's box. *Financial Services Review*. Vol.10. pp 129-144.

Şenol, E., Dinçer, H., & Timor, M. (2012). Artificial Neural Networks and Stock Selection Model Based on Fundamental and Technical Analysis Variables by Using Support Vector Machines. *Review of Economics & Finance*, Vol. 2, Issue 3, pp106 – 122.

Siqueira, E., Otuki, T. & Da Costa, N. (2012). Stock Return and Fundamental Variables: A Discriminant Analysis Approach. *Applied Mathematical Sciences*, Vol.6, Issue 115, ISSN: 5719 – 5733.

Sorensen, E. H., Miller, K.L. & Ooi, C.K. (2000). The Decision Tree Approach to Stock Selection. *The Journal of Portfolio Management.* Vol.27, Issue 1, pp 42 – 52.

Stancu Dumitra and Stancu Andrei-Tudor (2014). Predicting Company Performance by Discriminant Analysis. *Proceedings of the 8th International Management Conference*: *"Management Challenges for Sustainable Development",* November 6th – 7th, 2014, Bucharest, Romania.

Stevens, J. 91996). Applied Multivariate Statistics for the Social Sciences. In: Mahwah NJ, (3rd ed.) Lawrence Eribaum Associates.