Educational Research (ISSN: 2141-5161) Vol. 2(2) pp. 841-855 February 2011
Available online@ http://www.interesjournals.org/ER

Full Length Research Paper

Canonical correlation analysis -promotion bias scoring detector (a case study of American university of Nigeria)

Angus Okechukwu Unegbu[1], James J. Adefila[2]

[1]Assistant Professor of accounting, American University of Nigeria
[2]Senior lecturer, Department of Accountancy, University of Maiduguri, Borno state Nigeria

Accepted 17 February, 2011

The main focus of this paper is on Canonical Correlation Analysis as a bias scoring detector, using American University of Nigeria as a study case. In carrying out the research, we stipulated three null hypotheses. These hypotheses were stipulated after envisaging a problem that American University of Nigeria in a bid to keep with her value statement. The value statement stipulates that in all her activities, she will demonstrate the highest standards of integrity, transparency and academic honest but the University is faced with dwindling funding resulting in a cross road as to how to appraise and select candidates for promotion from numerous non-queuing applicants who have to secure highest weighted recommendations from various levels of promotion committees. The need for a bias free selection technique becomes very essential for adaption in analyzing applicants' scores from the various promotion committees. Scores of Candidates used in the research were obtained via secondary sources. Manual computations were first used to process the obtained scores and Computer SPSS Canonical Correlation Analysis were later used to process the results in order to test stated hypotheses. We found that Canonical Correlation Analysis has the capacity to detect bias scoring, overbearing score-weight influence and has the ability to discriminate between promotable and non-promotable candidates. We recommend its use to American University of Nigeria and other Universities of excellence. We also recommend for more research into the relevance of the statistical tool in the areas of Accounting and Business decision endeavors.

Keywords: Canonical Correlation, Promotion bias Score detector, American University of Nigeria

INTRODUCTION

American University of Nigeria (AUN) is a private University conceived in 2003 by a prominent Nigerian however in order to operate American styled institution, the University has to partner with American University Washington DC. AUN opened her doors to students in September 2005. According to AUN (2010), the mission of the university among others is providing an academic environment that promotes free thinking and research to benefit mankind. Continuing in her 'Values Statement' among others is that the 'University in all of her activities will demonstrate the highest standards of integrity, transparency, and academic honesty'. The University currently has three schools; School of Arts and Sciences, School of Business and

*Corresponding author Email: angusu1@gmail.com

Entrepreneurship and School of Information Technology & Communications. Each of these schools has many concentrations and programs requiring many faculties. Most of her faculties are employed from abroad and few from Nigerian localities. The number of faculties employed by the University has been on the increase since 2005 but she is yet to break-even based on admission policies which is in consonant with her vision and mission of 'offering international education standards & techniques at the door-steps of Africans'. In order to live to up to her 'Values Statement', the University Management set up various levels of Committees to appraise candidates (Faculties) for promotion bearing in mind that the University in all her activities do demonstrate the highest standards of integrity, transparency, and academic honest but within available resources. This inadequate resource (funds) has elicited the search for a deterministic tool suitable for selecting candidates for promotion free from bias among the numerous applicants from each school after sorting the recommendations of the various committees set up for appraising applicants for promotion, based on the highest weighted average recommendations from the numerous promotion committees.

Statement of the problem

American University of Nigeria in a bid to keep with her value statement which stipulates that in all her activities, she will demonstrate the highest standards of integrity, transparency and academic honest is faced with dwindling funding resulting in a cross road as to how to appraise and select candidates for promotion from numerous non-queuing applicants who have to secure highest weighted recommendations from various levels of promotion committees. The need for a bias free selection technique becomes very essential for adaption in analyzing applicants' scores from the various promotion committees. Canonical Correlation Analysis statistical technique is therefore considered as essential tool for addressing this problem.

Research Objectives

The main objective of this research is test the efficacy of Canonical Correlation Analysis as a relevant statistical tool for adaption in bias free promotion score processing and promotion bias scoring detector so as to ensure fairness, integrity, transparency and academic honest in analysis of applicants' scores and in reaching Faculties' promotion decisions.

Research Hypotheses

Ho_1: Canonical Correlation Analysis cannot detect bias scoring for any of the candidates from any of the named Promotion Committees with 90% confidence level.

$Ho_{2:}$ Canonical Correlation Analysis cannot detect significantly whether or not score-weights of each of the Promotion Assessors have over bearing influence on the promotability of candidates.

Ho_3. Canonical Correlation Analysis cannot at 90% level of certainty discriminate between candidates that have earned promotion scores and those that could not from various promotion committees of the University.

Significance of the study

The study highlights the adequacy of Canonical Correlation Analysis in decision making when faced with multivariate decision variables and parameters which must be taken into the overall considerations to enable optimization of scare resources within limiting constraints in order to arrive at unbiased decision especially in Universities' promotion exercises and American University of Nigeria in particular.

Limitations and scope of the study

This research is limited to the study of use of Canonical Correction Analysis as a bias scoring detector tool for selection of Faculties for promotion. This multivariate statistical tool which can be adapted for Faculties' promotion deterministic tool and bias scoring detector is tested using the data from American University of Nigeria for only one promotion excise because of exigencies of time and available resources for the research.

Review of related literature

ResearchConsultation.Com (2007) defined Canonical Correlation as a multivariate technique that allows you to assess the relationship between two sets of variables- the predictor set and the criterion variable set. Garson (2008) posits that a canonical correlation is the correlation of two canonical (latent) variables, one representing a set of independent variables, the other a set of dependent variables. Aaron, Sally and Steve (2005) asserted that canonical correlation is used to investigate the overall correlation between two sets of variables (***p'*** and ***q'***). The basic principle behind canonical correlation is determining how much variance in one set of variables is accounted for by the other set along one or more axes.

There are several measures of correlation that express the relationship between two or more variables. The standard Pearson product moment correlation coefficient (*r*) measures the extent to which two variables are related. *Multiple Regression* allows one to assess the relationship between a dependent variable and a set of independent variables. Multiple correspondences Analysis is useful for exploring the relationships between a set of categorical variables.

Canonical correlation is different from Multiple regression because whereas multiple regression is used for many-to-one relationships, canonical correlation is used for many-to-many relationships (Garson, 2008). Statsoft (2007) posits that Canonical Correlation is an additional procedure for assessing the relationship between variables. Specifically, this analysis allows us to investigate the relationship between *two sets* of variables. For example, an educational researcher may want to compute the (simultaneous) relationship between three measures of scholastic ability with five measures of success in school. Canonical correlation is also called a characteristic root. Wuensch (2009) explained that the goal of canonical correlation is to describe the relationships between two set of variables. Borga (2001) traced the origin of Canonical analysis to the work of Hotelling(1935, 1936). Hardoon ,Szedmak and Shawe-Taylor (2003) agreed to Borga(2003)'s assertion as to the origin of canonical correlation analysis. Hardoon ,Szedmak and Shawe-Taylor (2003) asserted that Canonical correlation analysis can be seen as the problem of finding basis vectors for two sets of variables such that the correlation between the projections of the variables onto these basis vectors are mutually maximized. Correlation analysis is dependent on the co-ordinate system in which the variables are described, so even if there is a very strong linear relationship between two sets of multidimensional variables, depending on the co-ordinate system used, this relationship might not be visible as a correlation. Canonical correlation analysis seeks a pair of linear transformations one for each of the sets of variables such that when the set of variables are transformed the corresponding co-ordinates are maximally correlated. Borga (2001), however defined canonical correlation analysis as the problem of finding two sets of basis vectors, one for x and the other for y, such that the correlations between the projections of the variables onto these basis vectors are mutually maximized.

Types of canonical correlation analysis

Generally, there are two main types of canonical correlation analysis. According to Mishra(2009), they are; ordinal canonical correlation coefficient and conventional canonical correlation analysis. The ordinal canonical correlation coefficient, $r(Z_1,Z_2)$, is the coefficient of correlation between two ordinal variables (Z_1 and Z_2), both of them being the composite (ordinal) ranking scores derived from two ordinal multidimensional data sets of ranking scores,X_1 and X_2, such that $r(Z_1,Z_2)$ is of the largest magnitude. It may be considered analogous to the conventional coefficient of canonical correlation in which the composite canonical variates (Y_1 and Y_2) are cardinally measured. It may be noted that while X_1 and X_2 are in themselves the ordinal variables, their transformation to cardinally measured canonical variates is problematic. Therefore, in such conditions, the ordinal coefficient of correlation (an analog of Spearman's rank correlation) would be a more appropriate measure of concordance between two sets of variables (that is, the ranking scores).

The Conventional Canonical Correlation Analysis: The conventional canonical correlation analysis (Hotelling, 1936) maximizes the squared (product moment) coefficient of correlation between two composite variates (Y_1 and Y_2) obtained as a linear combination of two sets of data, X_1 and X_2, on m1 and m_2 variables (respectively) each in n observations [n > max(m1, m_2)linearly independent cases]. It is a straightforward (multivariate) generalization of (Karl Pearson's product moment coefficient of correlation.

Key concepts and terms

Variables

These are what you have measured in your research (Aaron, Sally and Steve, 2005:1). According to

Garson (2008:2) a canonical variable, also called a variate, is a linear combination of a set of original variables in which the within-set correlation has been controlled (that is, the variance of each variable accounted for by other variables in the set has been removed). It is a form of latent variable. There are two canonical variables per canonical correlation (function). One is the dependent canonical variable, while the one for the independents may be called the covariate canonical variable.

Pooled R_c^2 (pooled canonical correlation) is the sum of the squares of all the canonical correlation coefficients, representing all the orthogonal dimensions in the solution by which the two sets of variables are related. Pooled R_c^2 is used to assess the extent to which one set of variables can be predicted or explained by the other set. (Aaron, Sally and Steve, 2005:1-2)

Eigenvalues

They reflect the proportion of variance in the canonical variate explained by the canonical correlation relating two sets of variables. According to Field (2005:590), eigenvalues are conceptually equivalent to the F-ratio in Analysis of Variance (ANOVA). The function of eigenvalues is to compare the obtained values with those we expect by chance alone. Field (2005:590-592) opined that there are four ways of obtaining the eigenvalues. These are;

a. Pillai-Bartlett trace (V): Pillai's trace is the sum of the proportion of explained variance on the discriminant functions. It is computed thus:

$$V = \sum_{i=1}^{s} \frac{\lambda_i}{1+\lambda_i}.$$

It is similar to the ratio of SS_M/SS_T

b. Hotelling's T^2: This is test statistic of sum of SS_M/SS_R for each of the variates. It is computed thus:

$$T = \sum_{i=1}^{s} \lambda_i$$

c. Wilk's Lamba (Λ): This is the product of the unexplained variance on each of the variates. π symbol is similar to the summation symbol (Σ). It therefore means that Wilk's lambda represents the ratio of error variance to total variance (SS_R/SS_T) for each variate. It is computed thus:

$$\Lambda = \prod_{i=1}^{s} \frac{1}{1+\lambda_i}$$

d. Roy's Largest root: This represents the proportion of explained variance to unexplained variance (SS_M/SS_R) for the first discriminant function. It is computed thus: Largest root = $\lambda_{l\arg est}$

Canonical weight

Also called the canonical function coefficient or the canonical coefficient: the standardized canonical weights are used to assess the relative importance of individual variables' contributions to a given canonical correlation. The canonical coefficients are the standardized weights in the linear equation of variables which creates the canonical variables. As such they are analogous to beta weights in regression analysis. The ratio of canonical weights is the ratio of the contribution of the variable to the given canonical correlation, controlling for other variables in the equation. There will be one canonical coefficient for each original variable in each of the two sets of variables, for each canonical correlation. Thus for the dependent set, if there are five variables and there are three canonical correlations (functions), there will be 15 canonical coefficients in three sets of five coefficients.

Canonical scores

Are the values on a canonical variable for a given case, based on the canonical coefficients for that variable. Canonical coefficients are multiplied by the standardized scores of the cases and summed to yield the canonical scores for each case in the analysis.

Structure correlation coefficients

Also called canonical factor loadings: a structure correlation is the correlation of a canonical variable with an original variable in its set. That is, it is the correlation of canonical variable scores for a given canonical variable with the standardized scores of an original input variable. The table of structure correlations is sometimes called the factor structure. The squared structure correlation indicates the contribution made by a given variable to the explanatory power of the canonical variate based on the set of variables to which it belongs. Alpert and Peterson (1972: 192) noted that canonical weights appear more suitable for prediction, while correlations or structure coefficients may better explain underlying (although interrelated) constructs.

Type of data for which it is meant

Canonical Correlation Analysis is useful for analysis of the relationship between two or more set of descriptive variables. Hardoon , Szedmak & Shawe-Taylor (2003) listed examples of its applicability such as; an educational researcher may want to compute the (simultaneous) relationship between three measures of scholastic ability with five measures of success in school. A sociologist may want to investigate the relationship between two predictors of social mobility based on interviews, with actual subsequent social mobility as measured by four different indicators. A medical researcher may want to study the relationship of various risk factors to the development of a group of symptoms. In all of these cases, the researcher is interested in the relationship between two sets of variables, and Canonical Correlation would be the appropriate method of analysis. According to ResearchConsultation.Com (2007) a canonical correlation is used when there are multiple continuous dependent and independent variables and the goal of your dissertation or thesis is to assess the relationship between these two sets of variables. Let's suppose your dissertation is concerned with a set of variables measuring attractiveness (physical beauty, warmth, kindness, and sex appeal) and a set of demographic characteristics (religious affiliation, charitable contributions, socioeconomic status, and income). Your analysis includes a canonical correlation and it reveals two sets of variable relationships. That is, there are two reliable ways these two sets of variables are related.

How it can be applied

Canonical Correlation Analysis can be applied using step by step method. It is easier however to apply Canonical Correlation Analysis with the use of Computer software such as SPSS. It is assessed on loading the data file via; Analyze to General Linear Model to Multivariate and keeping tag of Dependent and Independent variables.

Shortcomings or limitations

Wuensch (2009) posits that Canonical Correlation is subject to several limitations. It is mathematically elegant but difficult to interpret because solutions are not unique. Procedures that maximize correlation between canonical variate pairs do not necessarily lead to solutions that make logical sense. It is the canonical variates that are actually being interpreted and they are interpreted in pairs. A variate is interpreted by considering the pattern of variables that are highly correlated (loaded) with it. Variables in one set of the solution can be very sensitive to the identity of the variables in the other set; solutions are based upon correlation within and between sets, so a change in a variable in one set will likely alter the composition of the other set. There is no implication of causation in solutions. The pairings of canonical variates must be independent of all other pairs. Only linear relationships are appropriate. ResearchConsultation.Com (2007 specifically listed the limitations of Canonical correlation in dissertation research as follows; First, you have variables, then canonical variates, and canonical variate pairs. Variables are your measured dissertation variables (e.g. socioeconomic status, physical beauty, religious affiliation). Canonical variates are linear combinations of your variables, with one combination on the independent variable side (socioeconomic status and income), and one combination on the dependent side (physical beauty and sex appeal). These two combinations make up a pair of canonical variates. Of course, there may be more than one reliable pair of canonical variates. The biggest of these limitations is interpretability. Although mathematically elegant, canonical solutions are often un-interpretable. Furthermore, the rotation of canonical variates to improve interpretability is not a common practice in research, even though it is commonplace to do this for factor analysis and principle components analysis. Another problem using canonical correlation for research is that the algorithm used emphasizes the linear relationship between two sets of variables. If the relationship between variables is not linear, then using a canonical correlation for the analysis may miss some or most of the relationship between variables.

Usefulness and advantages

In the opinion of Joseph, Rolph, Ronald and William (1998) Canonical correlation analysis is a useful and powerful technique for exploring the relationships among multiple dependent and independent variables. The technique is primarily descriptive, although it may be used for predictive purposes.

Results obtained from a canonical analysis should suggest answers to questions concerning the number of ways in which the two sets of multiple variables are related, the strengths of the relationships, and the nature of the relationships defined. Canonical analysis enables the researcher to combine into a composite measure what otherwise might be an unmanageably large number of bivariate correlations between sets of variables. It is useful for identifying overall relationships between multiple independent and dependent variables, particularly when the data researcher has little a priori knowledge about relationships among the sets of variables. Essentially, the researcher can apply canonical correlation analysis to a set of variables, select those variables (both independent and dependent) that appear to be significantly related, and run subsequent canonical correlations with the more significant variables remaining, or perform individual regressions with these variables. One of the usefulness of canonical correlation analysis is as captured by Garson (2008) as he stated that canonical correlation is optimized such that the linear correlation between the two latent variables is maximized. Whereas multiple regressions are used for many-to-one relationships, canonical correlation is used for many-to-many relationships. There may be more than one such linear correlation relating the two sets of variables, with each such correlation representing a different dimension by which the independent set of variables is related to the dependent set. The purpose of canonical correlation is to explain the relation of the two sets of variables, not to model the individual variables. It is appropriate when the researcher desires to parsimoniously describe the number and nature of independent relationships that exist between the two sets (Stevens, 1996). Another advantage of canonical correlation analysis by the posits of Joseph, Rolph , Ronald and William (1998) is that it has gained acceptance in many fields and represents a useful tool for multivariate analysis, particularly as interest has spread to considering multiple dependent variables.

Its uniqueness

Garson (2008) stated that whereas multiple regression is used for many-to-one relationships, Canonical correlation is used for many-to-many relationships. There may be more than one such linear correlation relating the two sets of variables, with each such correlation representing a different dimension by which the independent set of variables is related to the dependent set. The purpose of canonical correlation is to explain the relation of the two sets of variables, not to model the individual variables. According to Joseph, Rolph, Ronald and William (1998) Canonical correlation places the fewest restrictions on the types of data on which it operates. Because the other statistical techniques impose more rigid restrictions, it is generally believed that the information obtained from them is of higher quality and may be presented in a more interpretable manner. For this reason, many researchers view canonical correlation as a last-ditch effort, to be used when all other higher-level techniques have been exhausted. But in situations with multiple dependent and independent variables, canonical correlation is the most appropriate and powerful multivariate technique.

How to interpret the result of canonical correlation analysis

There are known three methods of interpretation as stated by Pearson (2010). The three methods for interpretation are (1) canonical weights (standardized coefficients), (2) canonical loadings (structure correlations), and (3) canonical cross-loadings. The canonical function can be interpreted by the sign and the magnitude of the canonical weights assigned to each variable in its respective canonical variate. Variables with larger weights contribute more to the variates, and vice versa. Because canonical weights are derived to maximize the canonical correlations, they are subject to considerable instability from one sample to another. Also, weights may be distorted due to multicollinearity. Therefore, considerable caution is necessary if interpretation is based on canonical weights. Canonical loadings measure the correlation between the original observed variables and its canonical variate. They can be interpreted like factor loadings. Variables with larger loadings are more important in deriving the canonical variate. Whereas weights are more suitable for prediction, loadings are better at explaining underlying constructs. Canonical loadings are considered more valid and stable than weights. Canonical cross-loadings measure the correlation between the original observed variables and their opposite variate (i.e., independent variables correlate to the dependent variate, dependent variables correlate to the independent variate). They offer more direct interpretations by eliminating an intermediate step of conventional loadings.

Brief on AUN promotion procedure

According to AUN (2009), the Committees\relevant authorities that make input into a candidate's promotion application are; Weights

a. Dean of the School 7.5%
b. School Promotion Committee 7.5%
c. The Academic Vice President 10%
d. External Assessor/Reviewer 10%
e. University Wide Promotion Committee 15%
f. The Senate Committee 20%
g. President of the University 30%

Total 100%

The benchmark for promotion is securing a weighted average score of ≥65%age. The score rankings of the above listed authorities can be continently grouped into;

a. Internal Academic Assessors: 25% weights, comprising of scores from Deans, School Promotion Committees and Academic Vice President.
b. External Academic Assessors: 45% weights, made up scores from External Assessor, University Wide Promotion Committee, and The Senate Council.
c. The President of the University: 30% weights based on your academic records and contributions to the University.

Each of the Committee's point allocation will be based thus;

i. Teaching Effectiveness 35 – 45 %
ii. Scholarship (research & creative works 35 – 45 %
iii. Service to the University & to Community 20 – 25 %

Supporting documents for Teaching;

1. Peer evaluation
2. Student evaluation
3. Course Syllabi

Letters of recommendation highlighting teaching ability

4. Record of participation in teaching seminars, workshops, etc
5. Record of former students gainfully employed in related professions in the major
6. Evidence of innovative teaching methods and/or integration of technology in the delivery of course materials to students. This should include a self-assessment with supporting portfolio
7. Contributions to the development of new academic programs, course concentrations, and/or new courses
8. Sample of works by students
9. Commendations from students, Student's parents' or others in appreciation of candidate's efforts as a teacher.
10. Faculty awards for excellence in teaching.

Scholarship, Research and Creative Works

1. Terminal degrees/Professional qualifications
2. At least Five publications, three of which shall be journal articles
3. Computer Software and Program development
4. Television documentaries, film, video, or radio productions
5. Screenplays, videotapes etc
6. Publications, exhibition, or broadcast of photographic and graphic materials
7. Creative work in the areas of advertising, public relations, layout design, photography and graphics, visual arts etc.

Service to the University, Profession and Community;

1. Membership/leadership in departmental, school-wide and/or university-wide committees.
2. Exemplary departmental, school-wide and /or university-wide citizenship.
3. Advisement of student organization.
4. Planning and /or participation in workshops, conferences, seminars or presentations.
5. Recruitment of outside speakers to department-wide or campus-wide audiences.
6. Evidence of participation in university fundraises and/or grants writing activities.
7. Evidence of participation in mentoring or career counseling of students.
8. Expert testimony as part of the public record.
9. Board of directorships, advisory board membership etc

Expert views as public commentary on radio and television programs, newspapers, magazines, etc.

10. Membership in Civil Society organizations.

11. Community workshops, seminars and presentations.

12. Participate in consulting/advising activities.

13. Evidence of membership in professional organizations.

14. Evidence of intellectual leadership, eg editorial work and reviews for professional journals.

15. Evidence of service as external assessor or external examiner on examination committees.

It is the responsibility of each Applicant to prepare and present his/her dossier to each of the Dean of his/her school and the Academic Vice President on stipulated dates. The Academic Vice President will deliver copies of the dossier to all relevant Promotion authorities. At a stipulated date, the Deans, School Promotion Committees will submit their scores to the AVP (with advance copies to the President). Scores submitted to the Academic Vice President, including his own scores will be submitted to the President in sealed envelopes on a predetermined date. External Reviewer, University Wide Promotion Committee, and Senate Committee will directly submit to the President. The President and the Academic Vice President will meet to open the sealed submissions and have the scores first compiled in a raw form with the help of responsible officials in attendance and then weights them for final computations. The outcome is then used as bases for promotion recommendations to the University Board of Trustees for approval and ratification.

METHODOLOGY

The scores of each applicant from the various Promotion Committees were obtained from secondary sources. These data were subjected to Canonical Correlation analysis. The outcomes of Promotion Committees' report computed manually were tested with those generated from Computer SPSS Canonical Correlation analyses in order to test stipulated null hypotheses. The Conventional Canonical Analysis is employed. The conventional canonical correlation analysis (Hotelling, 1936) maximizes the squared (product moment) coefficient of correlation between two composite variates (Y_1 and Y_2) obtained as a linear combination of two sets of data, X_1 and X_2, on m1 and m_2 variables (respectively) each in n observations [n > max(m1, m_2)linearly independent cases]. It is a straightforward (multivariate) generalization of (Karl Pearson's product moment coefficient of correlation. By the work of Aaron, Sally, & Steve (2005) canonical correlations analysis is mathematical formed by first, a correlation matrix (**R**) is formed. **R** is the product of the inverse of the correlation matrix of ***q'*** (**R_{yy}**), a correlation matrix between ***q'*** and ***p'*** (**R_{yx}**), the inverse of correlation matrix of ***p'*** (**R_{xx}**), and the other correlation matrix between ***q'*** and ***p'*** (**R_{xy}**).

$$\mathbf{R} = \mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\mathbf{R}_{xx}^{-1}\mathbf{R}_{xy}$$

Canonical analysis proceeds by solving the above equation for eigenvalues and eigenvectors of the matrix **R**. Eigenvalues consolidate the variance of the matrix, redistributing the original variance into a few composite variates. Eigenvectors, transformed into coefficients, are used to combine the original variables into these composites. The eigenvalues are related to the canonical correlation by the following equation: $\lambda_i = r_{ci}^2$. That is, each eigenvalue equals the squared canonical correlation for each pair of canonical variates.

The significance of one or more canonical correlations is tested as a chi-square variable using the following formula:

$$\chi^2 = -\left[N - 1 - \left(\frac{k_x + k_y + 1}{2}\right)\right]\ln \Lambda_m$$

$$\Lambda_m = \prod_{i=1}^{m}(1-\lambda)$$

with,

- **N**= number of cases
- **k_x**= number of variables in ***p'***
- **k_y**= number of variables in ***q'***
- **df**= **(k_x)(k_y)**
- **m**= number of canonical correlations

Significant results indicate that the overlap in variability between variables in the two sets is significant; this is evidence of significance in the first canonical correlation. This process (of finding a canonical correlation and testing for significance) is then repeated with the first pair of variates removed to see if any of the other pairs are significant. Only pairs that test significant are interpreted. Canonical coefficients (also referred to as canonical weights) are analogous to the beta values in regression. One set of canonical coefficients is required for each set of variables for each canonical correlation. To facilitate comparisons, these values are usually reported for standardized variables (z transformed variables). The coefficients reflect differences in the contribution of the variables to the canonical correlation. Canonical correlation analysis is a useful and powerful technique for exploring the relationships among multiple dependent and independent variables that are unique in AUN promotion exercise. The result of Canonical Correlation is though mathematically elegant but it difficult to interpret because solutions are not unique to known AUN decision parameter.

Data presentation

Table 1: Applicants' ratings by Various Promotion Committees (Source: President's Office)

In order to maintain the sensitivity of the information, candidates' names are hereby omitted and the list does not represent any given real order.

A glance at the raw scores of Table 1 above shows that the order of promotion rankings would have been;

1st ranks: Applicant numbers 2 & 14.

Table 1. Raw Scores of Candidates

		S	C	O	R	E	S	
Candidates	**Dean**	**SPC**	**AVP**	**EAR**	**UWPC**	**TSC**	**POU**	**TOTAL**
1	60	80	50	60	70	70	40	430
2	80	60	60	50	70	70	80	470
3	80	60	50	40	70	60	50	410
4	80	70	60	40	40	60	90	440
5	95	90	85	70	50	40	30	460
6	20	40	70	65	70	75	80	420
7	80	60	60	50	60	50	40	400
8	60	60	70	50	70	60	60	430
9	50	70	60	75	65	60	50	430
10	35	40	50	50	65	80	80	400
11	90	80	60	70	40	40	80	460
12	65	80	75	50	60	50	50	430
13	40	30	45	40	75	80	80	390
14	70	90	80	70	60	50	50	470

Where;

Candidates: Applicants for Promotion
Dean: Dean of each School's Score of the candidates.
SPC: School Promotion Committees' Score of each candidate.
AVP: Academic Vice President's Score of each candidate.
EAR: External Assessor/Reviewer's score of each of the candidates.
UWPC: University Wide Promotion Committee's Score of each candidate.
TSC: The Senate Council's Score of the candidates.
POU: President of the University's score of the candidates.
TOTAL: Total Score earned by each candidate. Each of the scoring authority scores is over 100%.

2nd ranks: Applicant numbers 5 & 11.
3rd ranks: Applicant number 4.
4th ranks: Applicant numbers 1, 8, 9 & 12.
5th ranks: Applicant number 6.
6th ranks: Applicant number 3.
7th ranks: Applicant numbers 7 & 10.
8th ranks: Applicant number 13.

The rankings above based on the order of scores is far from the reality on applying weights to each of them as seen in manual processing of the scores in Table 2

Result of manual processing is that Candidates with Numbers 2, 4, 6, 10 and 13 who earned Total scores of ≥ 65% were recommended to the Board of Trustees for approval for promotion to the next ranks and these were dully ratified.

RESULTS AND INTERPRATIONS

Processing Procedure is: 1st Step: Capture the Candidates into two Groups of Promotable and Non-promotable. Also clearly show the score of these candidates from various groups of Internal Academic Assessors (IAA), External Academic Assessors (EAA) and President of the University (POU). The table will appear thus (Table 3);

Step 2: Code the variables into the Variable view via coding of the Candidates into two Statuses of promotable and non-promotable, while the independent variables such as scores from the appraising authorities and committees are entered. The data input view containing both the three groups of assessors and individual assessors will appear thus (Table 4);

.Step 3: The command line to give at the Data view are two;

One is; **Analyze** $\Rightarrow$ **General Linear Model** $\Rightarrow$ **Multivariate** and then key test parameters.

The second one is; **Analyze** $\Rightarrow$ **Classify** $\Rightarrow$ **Discriminant.**

The SPSS computer results

The results on keying the test parameters from SPSS will appear thus (table 5-7);

Interpretations of tables 5-7

Table 5, 6 and 7 classified applicants/candidates into two groups of promotable and non-promotable of

Table 2. Processed scores of the Candidates

Weights Candidates	7.5% Dean	S 7.5% SPC	C 10% AVP	O 10% EAR	R 15% UWPC	E 20% TSC	S 30% POU	Total (%)	Remark
1.	4.5	6	10	6	10.5	14	12	58	NP
2.	6	4.5	6	5	10.5	14	24	70	P
3.	6	4.5	5	4	10.5	12	15	57	NP
4.	6	5.25	6	4	6	12	27	66.25	P
5.	7.13	6.75	8.5	7	7.5	8	9	53.8	NP
6.	1.5	3	7	6.5	10.5	15	24	67.5	P
7.	6	4.5	6	5	9	10	12	52	NP
8.	4.5	4.5	7	5	10.5	12	18	61.5	NP
9.	3.75	5.25	6	7.5	9.75	12	15	59.25	NP
10.	2.63	3	5	5	9.75	16	24	65.38	P
11.	6.75	6	6	7	6	8	24	63.75	NP
12.	4.88	6	7.5	5	9	10	15	57.38	NP
13.	3	2.25	4.5	4	11.25	16	24	65	P
14.	5.25	6.75	8	7	9	10	15	61	NP

Where **P**= Promotable
NP = Not Promotable.

Table 3. Scores of Promotable and Non-promotable Candidates from each Group Assessors

	Internal Assessors: 25 % Dean	SPC	Academic AVP	External 45% EAR	Academic UWPC	Assessors: TSC	POU: 30% POU	Total : 100%
PROMOTABLE								
2	6	4.5	6	5	10.5	14	24	70
4	6	5.25	6	4	6	12	27	66.25
6	1.5	3	7	6.5	10.5	15	24	67.25
10	2.63	3	5	5	9.75	16	24	65.38
13	3	2.25	4.5	4	11.25	16	24	65
NON-PROMOTABLE								
1	4.5	6	10	6	10.5	14	12	58
3	6	4.5	5	4	10.5	12	15	57
5	7.13	6.75	8.5	7	7.5	8	9	53.8
7	6	4.5	6	5	9	10	12	52
8	4.5	4.5	7	5	10.5	12	18	61.5
9	3.75	5.25	6	7.5	9.75	12	15	59.25
11	6.75	6	6	7	6	8	24	63.75
12	4.88	6	7.5	5	9	10	15	57.38
14	5.25	6.75	8	7	9	10	15	61

5 and 9 respectively. The results as shown in tables 5, 6 and 7 indicate the Canonical Correlation Analysis status discriminatory ability of grouping Candidates into promotable and Non-promotable status. The result leads to the rejection of Null hypothesis Ho3 which states that Canonical Correlation Analysis cannot with 90% confidence level discriminate between promotable and non-promotable candidates based on their earned scores.

Interpretations of tables 8 and 9

The Multivariate tests as shown in table 8 indicate the effect of scores of the group and individual assessors both on status determination and bias impact on such status. The figure shows that the F-ratio values and Hypothesis values (ie computed values and critical table values) differences are very insignificant. This means that the chances of each

Table 4: Data for Canonical analysis from the three Grouped Assessors and Individual Assessors:

No.	Status.	IAA.	EAA.	POU.	DEAN	SPC	AVP	EAR	UWPC	TSC.	PS	Total
1	1.0	20.5	42.5	12.0	4.0	6.0	10.0	6.0	10.5	14.0	12.0	58.0
2	0.0	22.5	29.5	24.0	6.0	4.5	6.0	5.0	10.5	14.0	24.0	70.0
3	1.0	15.5	26.5	15.0	6.0	4.5	5.0	4.0	10.5	12.0	15.0	57.0
4	0.0	17.25	22.0	27.0	6.0	5.25	6.0	4.0	6.0	12.0	27.0	66.25
5	1.0	22.38	22.5	9.0	7.13	6.75	8.5	7.0	7.5	8.0	9.0	53.8
6	0.0	11.5	32.0	24.0	1.5	3.0	7.0	6.5	10.5	15.0	24.0	67.25
7	1.0	16.5	24.0	12.0	6.0	4.5	6.0	5.0	9.0	10.0	12.0	52.0
8	1.0	16.0	27.5	18.0	4.5	4.5	7.0	5.0	10.5	12.0	18.0	61.5
9	1.0	15.0	29.25	15.0	3.75	5.25	6.0	7.5	9.75	12.0	15.0	59.25
10	0.0	10.63	30.75	24.0	2.63	3.0	5.0	5.0	9.75	16.0	24.0	65.38
11	1.0	18.75	21.0	24.0	6.75	6.0	6.0	7.0	6.0	8.0	24.0	63.75
12	1.0	18.38	24.0	15.0	4.88	6.0	7.5	5.0	9.0	10.0	15.0	57.38
13	0.0	9.75	31.25	24.0	3.0	2.25	4.5	4.0	11.25	16.0	24.0	65.0
14	1.0	20.0	26.0	15.0	5.25	6.75	8.0	7.0	9.0	10.0	15.0	61.0

Note: ‘No’ represents each of the named applicants.
Status: Those that are promotable are coded ‘0’ and those that are not are coded ‘1’.
he scores from each of the assessors are shown against each applicant

Table 5. Prior Probabilities for Groups

Promotability	Prior	Cases Used in Analysis	
		Unweighted	Weighted
Promotable	.500	5	5.000
Non Promotable	.500	9	9.000
Total	1.000	14	14.000

Table 6. Classification Results[a]

		Promotability	Predicted Group Membership		Total
			Promotable	Non Promotable	
Original	Count	Promotable	5	0	5
		Non Promotable	0	9	9
	%	Promotable	100.0	.0	100.0
		Non Promotable	.0	100.0	100.0

a. 100.0% of original grouped cases correctly classified.

candidate’s status determination resulting from scores across the assessors and those that might result from bias scoring are very insignificant. This result is further confirmed by the significance level of Pillar’s trace of 0.041, Wilk’s Lambda of 0.041, Hotelling’s trace of 0.041 and Roy’s Largest Root of 0.041. Therefore, since all of them showed that $p<0.05$, it means that there is no between-status differences in the scores between assessors of both group and individuals, thereby leading to the rejection of Null hypothesis (Ho_1) which states that Canonical Correlation Analysis cannot detect bias

Table 7. Between-Subjects Factors

		Value Label	N
Promotability	0	Promotable	5
	1	Non Promotable	9

Table 8. Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	1.000	1.628E3[a]	10.000	3.000	.000
	Wilks' Lambda	.000	.628E3[a]	10.000	3.000	.000
	Hotelling's Trace	5.425E3	1.628E3[a]	10.000	3.000	.000
	Roy's Largest Root	5.425E3	1.628E3[a]	10.000	3.000	.000
Status	Pillai's Trace	.971	10.078[a]	10.000	3.000	.041
	Wilks' Lambda	.029	10.078[a]	10.000	3.000	.041
	Hotelling's Trace	33.592	10.078[a]	10.000	3.000	.041
	Roy's Largest Root	33.592	10.078[a]	10.000	3.000	.041

a. Exact statistic
b. Design: Intercept + Status

scoring for any of the candidates from any of the named Promotion Committees with 90% confidence level. We therefore conclude that although there is no significant bias in the scores given to Candidates between-assessors but the scores of each assessor had a significant effect on the determination of each Candidate Status. The nature and extent of the effects cannot however be determined from the above figure. The test of Between-Subject Effects is shown in table 9. From the result, the overall extent of the effects of individual assessors' scores as compared with the candidates total scores in Status determination is very insignificant as $p=0.135>0.05$ while Total score of each Candidate is highly significant as $p=0.0001<0.05$.

Interpretations of tables 10 and 11

Overbearing score weight influence test hypothesis is aimed at detecting across the individual assessors' mark allocations and weights assigned to each rather than between the scores earned from individual assessors as compared with the total scores earned by each Candidate. From table 10, Levene's test should be non-significant for all the scores from the group and individual assessors if the assumption of homogeneity of variance has been met. The result as shown in table 11 indicates that except scores from Internal Assessors as evidenced from the Deans' Scores, the test shows that homogeneity of variance has been met. From table 11, the parameter estimates shows the weight assigned to each assessor or group of assessors and their influence on determination of Candidates' Status. The result indicates that;

a. For Group Assessors - Internal Assessors with $p=0.096$, External Academic Assessors with $p=0.526$ and The President's Assessment with $p=0.0001$, shows that except that of the President, the weight assigned to scores of other two are group assessors are insignificant.

b. For the Individual Assessors, the score weights assigned to School Promotion Committee,

Table 9. Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Individual Scores	3.506[a]	1	3.506	2.562	.135
	Total score	237.139[b]	1	237.139	22.480	.000
Intercept	Individual Scores	378.006	1	378.006	276.216	.000
	Total score	50193.219	1	50193.219	4.758E3	.000
Status	Individual Scores	3.506	1	3.506	2.562	.135
	Total score	237.139	1	237.139	22.480	.000
Individual A	Individual Scores	.000	0	.	.	.
	Total score	.000	0	.	.	.
Status * Individual A	Individual Scores	.000	0	.	.	.
	Total score	.000	0	.	.	.
Error	Individual Scores	16.422	12	1.369		
	Total score	126.584	12	10.549		
Total	Individual Scores	454.500	14			
	Total score	52892.949	14			
Corrected Total	Individual Scores	19.929	13			
	Total score	363.724	13			

Table 10. Levene's Test of Equality of Error Variancesa

	F	df1	df2	Sig.
Internal Academic Assessors	6.752	1	12	.023
External Academic Assessors	.276	1	12	.609
President's Assessment	1.360	1	12	.266
Dean Score	4.954	1	12	.046
School Promotion Committee	.946	1	12	.350
Academic Vice President	1.319	1	12	.273
External Assessors Review	1.222	1	12	.291
University Wide Promotion Committee	.316	1	12	.584
The Senate Committee	.425	1	12	.527
President Score	1.360	1	12	.266
Total Score	1.633	1	12	.225

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Status

The Senate Committee and the President were found very significant while those of Dean, Academic Vice-President, External Reviewer and University Wide Promotion Committee were found to be very insignificant.

The results lead us to reject the Null hypothesis

Table 11. Parameter Estimates

Dependent Variable	Parameter	B	Std. Error	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Internal Academic Assessors	Intercept	18.112	1.253	14.459	.000	15.383	20.842
	[Status=.0000]	-3.786	2.096	-1.806	.096	-8.353	.781
	[Status=1.0000]	0[a]	.	.	.	.	.
External Academic Assessors	Intercept	27.028	1.894	14.271	.000	22.901	31.154
	[Status=.0000]	2.072	3.169	.654	.526	-4.833	8.977
	[Status=1.0000]	0[a]	.	.	.	.	.
President's Assessment	Intercept	15.000	1.183	12.677	.000	12.422	17.578
	[Status=.0000]	9.600	1.980	4.849	.000	5.286	13.914
	[Status=1.0000]	0[a]	.	.	.	.	.
Dean Score	Intercept	5.362	.512	10.478	.000	4.247	6.477
	[Status=.0000]	-1.536	.856	-1.794	.098	-3.402	.330
	[Status=1.0000]	0[a]	.	.	.	.	.
School Promotion Committee	Intercept	5.583	.346	16.124	.000	4.829	6.338
	[Status=.0000]	-1.983	.579	-3.423	.005	-3.246	-.721
	[Status=1.0000]	0[a]	.	.	.	.	.
Academic Vice President	Intercept	7.111	.463	15.347	.000	6.102	8.121
	[Status=.0000]	-1.411	.775	-1.820	.094	-3.100	.278
	[Status=1.0000]	0[a]	.	.	.	.	.
External Assessors Review	Intercept	5.944	.390	15.244	.000	5.095	6.794
	[Status=.0000]	-1.044	.653	-1.601	.135	-2.466	.377
	[Status=1.0000]	0[a]	.	.	.	.	.
University Wide Promotion Committee	Intercept	9.083	.576	15.771	.000	7.828	10.338
	[Status=.0000]	.517	.964	.536	.602	-1.583	2.616
	[Status=1.0000]	0[a]	.	.	.	.	.
The Senate Committee	Intercept	10.667	.632	16.865	.000	9.289	12.045
	[Status=.0000]	3.933	1.058	3.717	.003	1.627	6.239
	[Status=1.0000]	0[a]	.	.	.	.	.
President Score	Intercept	15.000	1.183	12.677	.000	12.422	17.578
	[Status=.0000]	9.600	1.980	4.849	.000	5.286	13.914
	[Status=1.0000]	0[a]	.	.	.	.	.
Total Score	Intercept	58.187	1.083	53.746	.000	55.828	60.546
	[Status=.0000]	8.589	1.812	4.741	.000	4.642	12.536
	[Status=1.0000]	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

(Ho_2) which states that Canonical Correlation Analysis cannot detect significantly whether or not score-weights of each of the promotion assessors has overbearing influence on the promotability of candidates.

DISCUSSIONS, CONCLUSION AND RECOMMENDATIONS

American University of Nigeria in a bid to keep with her value statement which stipulates that in all her activities, she will demonstrate the highest standards of integrity, transparency and academic honest will see Canonical Correlation Analysis statistical technique as a relevant tool for analyzing her promotion marks in order to see and eliminate bias scoring and overbearing influence of score-weights assigned to the promotion Assessors. A glance at Table 1 shows the order of promotable rankings but application of Canonical Correlation Analysis results produced different ranking of candidates. To validate the robustness of Canonical Correlation Analysis on multiple dependent and independent variables that akin to AUN promotion exercise, hypothesis three which states that CCA cannot at 90% level of certainty discriminate between candidates that have earned promotion scores and those that could not from various promotion Committees of the University were postulated and tested. The results as shown in table 5 -7 indicate 100% validation of the alternative hypothesis. The adoption of Canonical Correlation Analysis on the promotion processing exercise of AUN will be interesting and challenging but more time will be required to ensure proper understanding of its application. Canonical Correlation Analysis alone cannot be used to generate promotion outcomes for American University of Nigeria but as an excellent check for bias scoring of Candidates. Another outcome of the research shows that the weight allotted to the President for promotion of a candidate has overbearing influence on the promotability of an applicant.

We recommend this statistical tool for other Universities of Excellence in processing their promotion exercises. We also ask for more academic researches into the uniqueness and relevance of Canonical Correlation Analysis in Entrepreneurship development, Accounting and business related areas.

REFERENCES

Aaron F, Sally C, Steve S (2005). Canonical Correlation & Principal Components Analysis.13/03/2010 from http://userwww.sfsu.edu/efc/classes/bio1710/pca/CCandPCA2.htm

AUN (2009). American University of Nigeria Criteria for Promotions, Policy Guidelines. P5 -6

AUN (2010). American University of Nigeria Admissions Guide

Alpert, Mark I, Peterson, Robert A (1972): On the interpretation of canonical analysis.J. Market. Res. 9(2), 187-192.

Borga M (2001): Canonical Correlation-*a Tutorial* 15/02/2010 from http://www.liu.se/-magnus/cca

Clark D (1975): Understanding Canonical Correlation Analysis 22/02/2010 from qmrg.org.uk/files/2008/12/3-understanding-canonical-correlation-analysis1.pdf

Field A (2005): Discovering Statistics Using SPSS Sage Publications Ltd London. P590 -592

Garson GD (2008): Canonical Correlation. Stat-notes of North Carolina State University

16/02/2010 from http://faculty.chass.ncsu.edu/garson/PA765/canonic.htm

Hardoon RD, Szedmak S, Shawe-Taylor J (2003): Canonical correlation analysis; An overview with application to learning methods StatSoft Electronic Statistics Textbook 17/02/2010 from http://www.statsoft.com/textbook/canonical-analysis

Joseph FH Jr, Roph EA, Ronald LT, William CB (1998): Multivariate Data Analysis, 5th edition.Prentice Hall, Inc. Pp. 10-300

ResearchConsultation.Com (2007): Canonical Correlation in Dissertation & Thesis Research.

13/02/2010 from http://www.researchconsultation.com/canonical-correlation-statistics.asp

Mishra SK (2009): A note on the ordinal canonical correlation analysis of two sets of ranking Scores21/02/2010 from http://mpra.ub.uni_muechen.de/12796/1/MPRA_paper_12796.pdf

Pearson K (2009): *Multivariate Data Analysis*,Pearson Prentice Hall Publishing p185

Stevens J (1996). *Applied multivariate statistics for the social sciences,* 3rd edn.

Mahwah NJ: Erlbaum. P301.

Wuensch LK (2009): Canonical correlation.17/02/2010 from core.ecu.edu/psyc/wuenschk/MV/Canonical/Canonical.doc

.