The impact of morning shows and celebrity endorsement on the brands image using Multiple Correspondence Analysis

Dr Junaid Sageer Siddiqu[1], Dr Bushra Shamshad[2], Zara Omar[3]
[1]Professor, Department of Statistics, University of Karachi
[2]Assistant Professor, Department of Statistics, University of Karachi
[3]Lecturer, Department of Statistics, University of Karachi

Abstract
Now a days brands are having a heavy impact on the consumer and it causes a huge change in the purchase of products in the market. The impact of clothing and fashion brands can been seen in through the social media and in the peoples life. The celebrities and fashion brand ambassadors are making the images of the fashion and different clothing styles in the market. The thought behind this research is to explore the impact of morning shows on the clothing and fashion brands of the viewers. The purposes of the present study is (1) to find a linkage between morning show viewers having impact of brand endorsement (2) to empirically test the impact theory by examining the effect of the morning shows and brands image. Data is collected using an online survey, of the viewers of morning shows, the sample of 250 observation is used to test the factors. Correspondence analysis is an exploratory data analysis technique for the graphical display of contingency tables and multivariate categorical data. It is a useful method to find the association among bivariate and multivariate multidimensional variables. Results of the correspondence analysis shows that the morning shows have a positive and strong impact on the brand image and purchase of the fashion products. As predicted, brand endorsement is positively associated with morning shows and celebrity images.

Key words:
Correspondence analysis, Cronbach Alpha, VAF, Correspondence plots, brand endorsement, morning show, fashion brands, brand ambassador.

1. Introduction

Today a successful brand is one of the most important assets to many businesses. A brand is a unique element that make the products of an organization different from the competitors, also enhance the value of the product. Therefore, branding is the process of using a name, symbol, design, and experience to differentiate goods by providing distinct images, associations, and experiences. A consistent image, positive associations, and favorable attitudes formed from memorable experiences are essential in building a strong brand image. In the current situation, the morning show are performing the linkage between consumers and brands. It benefits both consumer and brands in a variety of ways. It would be easy for consumer to select the brand which suits them also for brands to convey its main theme for the season.

The morning show are playing an important role in it. It make the clear brand identity which helps marketers successfully differentiate their offerings from their competitors. Second, it successful branding helps firms reduce advertising costs by increasing awareness of the brand name. Third, morning show helps in becoming a leader among the competitors in the same product category by making its popular with in a show.
Morning shows also are providing consumers with three major benefits: risk reduction, information efficiency, and self-expression, because the celebrities have a great impact on the consumers and it helps consumers decrease the chance of choosing a product that may not perform well or meet their expectations. Thus, risk reduction is accomplished by providing assurance of consistent quality. Branding may help consumers recognize and become aware of offerings in a specific product category, which helps them efficiently categorize vast amounts of information available about the product.
For this purpose, the multiple correspondence analysis is used, which is an exploratory technique designed to analyses simple two-way and multi-way tables containing some measure of correspondence between the rows and columns. The results provide information which is similar in nature to those produced by Factor Analysis techniques, and they allow one to explore the structure of categorical variables included in the table. The most common kind of table of this type is the two or three-way frequency cross-tabulation table. Correspondence analysis (CA) may be defined as a special case of principal components analysis (PCA)

2. Literature review

The history of the development of multiple correspondence analysis started in the early 1960s with the work of Jean-Paul Benz´ecri and his team, who proposed the geometrical/graphical framework of the technique. Although one may also accredit Guttman (1941) as the founder of the technique in all, but its name, because his contributions were analogous in mathematical structure to what we now know as multiple correspondence analysis. Much of the pre-history of multiple correspondence

Manuscript received April 5, 2019
Manuscript revised April 20, 2019

analysis has been recorded by Jan de Leeuw at the University of California, Los Angeles; see, for example, De Leeuw (1983, 1988, 2006, 2014). Through his supervision, de Leeuw led a team of influential Dutch correspondence analysts who obtained their PhDs at the University of Leiden during the 1980s. These include Willem Heiser,who obtained his PhD in 1981,Piete r Kroonen berg in 1983,Jacqueline Meulman in 1986, Peter van der Heijden in 1987 and Jan van Rijckevorsel in 1987.

Multiple correspondence analysis is often viewed as a categorical analogy of principal component analysis and as a form of non-linear principal component analysis; see Pearson (1901, 1906) and Hotelling (1933). It is also viewed as a form of metric multidimensional scaling (De Leeuw, 1973; Hoffman and De Leeuw, 1991, 1992; Hoffman et al., 1994). Therefore, multiple correspondence analysis has been invented and reinvented many times under many varying frameworks over the years, but all with very similar goals. Recently, Le Roux and Rouanet (2005) proposed referring to the general approach as multivariate statistics, initiated by Benz´ecri with correspondence analysis, as geometric data analysis. Due to varying languages and varying data analytic visions across the globe (some refer to the matrices being analyzed in terms of objects × variables, some refer to variables ×variables), there are a wide variety of names for what turns out to be the same method. In the United States, it is often referred to as optimal scaling, optimal scoring and appropriate scoring; in Canada and Japan, it is referred to as dual scaling. In the Netherlands, multiple correspondence analysis is known as homogeneity analysis; in France and Italy, it is known by the name multiple correspondence analysis. This technique is also known as scalogram analysis in some parts of the world. Throughout this chapter, we refer to the technique as, simply, multiple correspondence analysis. The French phrase Analyse des Correspondences Multiples, which translates to ‘the analysis of multiple correspondences’, or ‘multiple correspondence analysis’ appears for the very first time in Lebart et al. (1977). Multiple correspondence analysis has been adopted by many disciplines to analyze large survey data as a highly informative and intuitive method for graphically depicting the association that exists between two or more categorical variables. This has largely been a result of the influential works of the researchers in Europe, including Greenacre (1984), Lebart et al. (1984), Gifi (1981), and De Leeuw (1973). Its developmental links with researchers from Japan is well understood along with the significant contributions of Nishisato (1980), Hayashi (1950) and Tanaka (1978). One of the hallmarks of correspondence analysis is the many ways in which one can derive the basic simultaneous equations. These equations are related to the Pearson’s chisquared statistic (Tenenhaus and Young, 1985; Van Rijckevorsel and De Leeuw, 1988a; Gifi, 1990; Nishisato, 1996; De Leeuw et al., 1999) and also to the different methods of decomposition/quantification. With correspondence analysis also referred to as dual scaling, the use of the word dual refers to the symmetrical manner in which the rows and columns of the data are analysed. This symmetry, or dual consideration of the two sets of categories, is an important mathematical property underlying the method (Nishisato, 1980, 1982, 1994; Tenenhaus, 1982; Nishisato and Nishisato, 1994; Van de Velden, 2000; Nishisato et al., 2002). It is interesting to note that the name dual scaling was first coined by Shizuhiko Nishisato in 1976, during a symposium held in France (Nishisato 2014).

The analysis of the association between the variables of a two-way contingency table may be considered a special case of multiple correspondence analysis. In practice, any two way contingency table can be obtained from a multi-way table by considering the product of the indicator matrix of one variable with the indicator matrix of another variable; refer to Section 10.2.2 for additional details. While the analysis of a two-way contingency table with a variable × variable structure has been widely discussed in the previous chapters, in this chapter, we focus on the correspondence analysis of a matrix subjects/individuals using categorical variables. For the sake of brevity,despitetherelevanceofstabilityandmodellingstudies formultiple correspondence analysis -- discussed by Grizzle (1969), Meulman (1984), Goodman (1986), Markus (1994), Linting et al. (2007), Van der Heijden et al. (1989), Van der Heijden and De Leeuw (1985), Lauro and De Carli (1982), Lauro and Siciliano (1988), Moritz Varnier et al. (2009)andMichailidisandDeLeeuw(2000)—and of the missing data problem (seeMeulman (1982) and Van Buuren and Van Rijckevorsel (1992)), we shall confine our attention to the treatment of multiple ordered categorical variables (Meulman et al., 2004; Lombardo and Beh, 2010; Lombardo and Meulman, 2010).

3. Multiple Correspondence Analysis

Correspondence analysis has the definition and statements in many books include [i]Greenacre (1984) and Lebart et ai (1984), which are famous. The basic technique is discussed in many publications such as , [ii]Greenacre (1981), Hoffman & Franke (1986), Wellar & Romney (1990), Benzecri (1992), Andersen (1994, 1997) and Everitt (1997). Recently mathematical presentation of correspondence analysis is also shown with the analysis of real life data. Goodman (1996) devised a single method which shows the equivalence of the methods of Pearson (1913), the work of G. U. Yule, and Fisher (1940). Goodman (1996) also generalized simple correspondence analysis with his RC association analysis of Goodman (1986).

3.1 Notations and Definition

Consider an $k \times m$ two-way contingency table, N, where the (k, m)'th cell entry is given by $n_{km}\ for\ k = 1,2,3 \dots . K, m = 1,2,3 \dots M.$ Let the grand total of N be n and the probability, or correspondence, matrix be Q that the (k,m)'th cell entry $q_{km} = \frac{n_{km}}{n}$ and

$$\sum_{k=1}^{K} \sum_{m=1}^{M} q_{km} = 1.$$

Define the k'th row marginal probability and m'th column marginal probability as

$$\sum_{k=1}^{K} q_{k.} = \sum_{m=1}^{M} q_{.m} = 1.$$

These marginal probabilities defines the mass values. Let $D_k\ and\ D_m$ be the diagonal matrix whose elements are the row mass and column mass respectively.

3.1.1 Inertia

The chi square is the measure of dependencies using variation from contingency the data obtained double total then relative variance between row and column will not change, in correspondence analysis $\frac{\chi^2}{n}$ is the value defines the total variance measure. This term is called inertia and each axis has a proportion called principal inertia. This is the reason correspondence analysis is dependent on the correspondence matrix Q not on the actual data.

3.1.2 Profiles

If the comparison is needed of the row and column values which is k and m. the probability of greater frequency values will be large. Profile calculation done on the row profiles and column profile of the contingency table.

3.1.3 Singular Value Decomposition

The correspondence analysis is a technique of multivariate which is used to calculate the scores describes difference of categories. The association or relationship between categories can be measure using row scores and column scores. Contingency table has rows and column categories of variables and correspondence can measure the association of variables in row and column based on the model of independence.

$$q_{km} = q_{k.}q_{.m} \tag{3.1}$$

In the case of multivariate data, variables are partially dependent which needs the measure of dependencies and model becomes:

$$q_{km} = \beta_{km} q_{k.}q_{.m} \tag{3.2}$$

For the sake of case simplification if $\beta_{km} \neq 1$, then the dependencies can be measure and if it is $\beta_{km} = 1$, then it is the case of independent variables.
The calculation of β_{km} is also possible and the formulation is called:

$$\beta_{km} = \frac{q_{km}}{q_{k.}q_{.m}} \tag{3.3}$$

Pearson ration and it was named by Goodman (1996). Darroch (1974) proposed the additive measure of the departure from independence:

$$\frac{q_{km}}{q_{k.}q_{.m}} = \nu_k + \omega_m$$

For all k and m , and for some $\{\nu\ \ \}\ and\ \{\omega\ \ \}.$
To determine the scoring of the rows and columns and the strength of the association between them, the Pearson ratios can be partitioned using the method of singular value decomposition:

$$\beta_{km} = \sum_{i=1}^{\acute{I}} a_{ki} \lambda_i\, b_{mi} \tag{3.4}$$

where,

$$\beta_{km} = \sum_{k=1}^{K} q_{k.} a_{ki} a_{k\acute{i}} = \begin{cases} 1, i = \acute{i} \\ 0, i \neq \acute{i} \end{cases} \tag{3.5}$$

and

$$\beta_{km} = \sum_{m=1}^{M} q_{.m} b_{mi} b_{m\acute{i}} = \begin{cases} 1, i = \acute{i} \\ 0, i \neq \acute{i} \end{cases} \tag{3.6}$$

And $\acute{I} = \min(k, m) - 1$

Consider the right hand side of equation (3.4). The set of values $\{a_{ku};\ k = 1, 2, \dots, K\}$ is the $u'th$ left generalised basic vector and is associated with the row categories. Similarly, the set of values $\{b_{mv}; m = 1, 2, \dots, M\}$ is the $v'th$ right generalised basic vector and is associated with the column categories. The generalized basic vectors are also referred to as singular vectors. The elements of $\{\lambda_0, \lambda_1, \lambda_2, \lambda_3 \dots . . \lambda_i\}$ are real and positive and are the first Í generalized basic values or singular values and are arranged in descending order so that:

$$1 = \lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \lambda_i \geq 0 \tag{3.7}$$

These generalized basic values can also be calculated by:

$$\lambda_i = \sum_{k=1}^{K} \sum_{m=1}^{M} q_{km} a_{ki}\, b_{mi} \tag{3.8}$$

The ordinal nature of the singular values of (3.7) show that while the first value is unity, or trivial, they have a minimum value of zero. The matrices of the singular vectors of A and B also contain a trivial solution to the problem as the set of values $\{a_{ko}\}and\ \{b_{mo}\}$ are equal to one.
In the terms of matrix notation the correspondence analysis can be write as:

$$D_K^{-1} Q D_M^{-1} = A D_\lambda B^T \tag{3.9}$$

Where,

$$A^T D_K A = I \qquad (3.10)$$

$$B^T D_M B = I \qquad (3.11)$$

where I denotes an identity matrix while D_K and D_M are as defined in notation heading. The matrix A contains the first $\acute{I}$ sets of row scores, referred to as left generalized basic vectors, and has dimension $K \times \acute{I}$ .Similarly, B contains the first $\acute{I}$ set of column scores called right generalized basic vectors and has dimension $M \times \acute{I}$. The matrix of singular values, D_λ is diagonal:

To remove this minor solution consider again equation (3.4). It becomes:

$$\beta_{km} = 1 + \sum_{i=1}^{\acute{I}} a_{ki} \lambda_i \, b_{mi} \qquad (3.12)$$

Goodman (1996) refers to $\beta_{km} - 1$ as the Pearson contingencies under model (2.1) for $k = 1, 2, \dots, K$ and $m = 1, 2, \dots, M$.

Therefore, (3.3) becomes:

$$\beta_{km} = \frac{q_{km}}{q_{k.} q_{.m}} = 1 + \sum_{i=1}^{\acute{I}} a_{ki} \lambda_i \, b_{mi}$$

or

$$\frac{q_{km} - q_{k.} q_{.m}}{q_{k.} q_{.m}} = \sum_{i=1}^{\acute{I}} a_{ki} \lambda_i \, b_{mi} \qquad (3.13)$$

4. Data analysis

Data is collected by online survey, the sample of 250 observations are used to find the impact of morning show and celebrity endorsement on the sale of branded cloths in the market of Pakistan. The research is based on viewers of morning shows who has impact of it in the clothing and fashion products which discussed in shows. Results from sample indicates that the impact of fashion brands and clothing is influences by the morning shows and celebrity endorsements.

Table 1: Correspondence table of the favorite morning shows and branded cloths

Correspondence Table				
Favorite morning show	**Clothing Brands**			
	Alkaram	**J.**	**Maria B**	**Active Margin**
Good Morning Pak	5	33	4	42
Jago Pakistan Jago	22	16	18	56
Subha Sawera Samma Ka sath	26	17	27	70
The Morning Show	7	42	4	53
Active Margin	60	108	53	221

Table 1 consist of the data of people watch mornings shows also believe on the branded cloths purchase because of the morning shows and celebrity endorsement. Categories in the rows and columns of Table 1 consist an ordinal structure, correspondence analysis is conducted on the data. The correspondence plot is given by Figure 1 and consists of the two dimension first is location and second is dispersion.

Firstly the dimension 1, has the inertia value of 0.276, accounts for 99.2% of the total inertia, while the second dimension 2 has the inertia value of 0.002, accounts for only 8%. This variations shows that the majority of the variation in the morning shows due to the difference in their mean values and their spread across the different brands of cloths.

Table 2: Profile table of the favorite morning shows and branded cloths

Dimension	**Singular Value**	**Inertia**	**Chi Square**	**Sig.**	**Proportion of Inertia**	
					Accounted for	**Cumulative**
1	0.525	0.276			0.992	0.992
2	0.046	0.002			0.008	1
Total		0.278	61.368	.000a	1	1
a. 16 degrees of freedom						

Table 2 shows the impact of morning shows on the purchase of branded cloths. It has the Pearson chi-squared statistic 61.33 and has a p-value that is less than 0.000, which shows the association between variables. It is a statistically significant association between the morning shows and branded cloths purchase. The dimension 1 and 2 has the value of total inertia 0.278. The first category is showing impact on the other factors which is very clear. The first eigenvalue is showing an important role in the whole analysis as it is showing strong association among the variables in the analysis of the variables. The table shows that second dimension can be seen to contribute equally to the dispersion component.

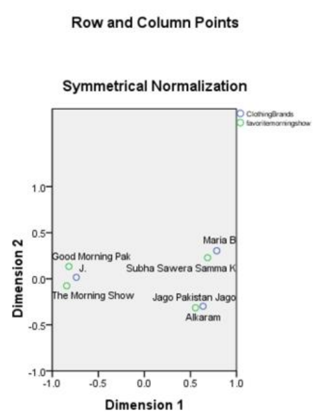

Fig. 1 Correspondence plot for Ordinal data set in Table1 (Dimension 1 and 2)

For the row and column profiles, only the location component is significant, thus a one-dimensional ordered correspondence plot is adequate to display the rows and columns. The dimension 1 being the first principal axis, described by the location component. It was shown that from Figure 1 that the mornings show views and brand image are having similar profiles. From Figure 1, it is showing that they are similar in terms of the location component and the dispersion component (with a slight difference in terms of dispersion). The morning show Jago Pakistan Jago and Alkaram are similar in terms of their mean value and their spread. The row and column contributions to the components are reflected by the distances of the profiles from the configurations centroid. Categories in row jago Pakistan Jago and Good Morning Pakistan are unlike in terms of the centroid along the location axis, it is expected that these profiles contribute to the location inertia.

Table 3: Component Values for Row Profiles of the favorite morning shows and branded cloths

Row Profiles				
Favorite morning show	Clothing Brands			
	Alkaram	J.	Maria B	Active Margin
Good Morning Pak	0.119	0.786	0.095	**1**
Jago Pakistan Jago	0.393	0.286	0.321	**1**
Subha Sawera Samma Ka sath	0.371	0.243	0.386	**1**
The Morning Show	0.132	0.792	0.075	**1**
Mass	**0.271**	**0.489**	**0.24**	

The above table shows the row profiles of the variable branded cloths. For the row profiles, which has inertia of J. is the highest .489, represents 48.9% of the total inertia, while the second highest has inertia of Alkaram is 0.271, represents 27.1%. From Table 3 it can be seen that J. has contribute 48.90% to this axis and it has 50% of the total inertia of the row profile. That is, nearly all of the variation in the values of the branded cloths category is due to the significant difference among J. Alkaram and Maria B. Other brands are situated close to the centroid along the location axis, these profiles account for very little of that component. They contribute to the remaining 50% of the location inertia.

Table 4: Component Values for Colum Profiles of the favorite morning shows and branded cloths

Column Profiles				
Favorite morning show	Clothing Brands			
	Alkaram	J.	Maria B	Mass
Good Morning Pak	0.083	0.306	0.075	0.19
Jago Pakistan Jago	0.433	0.157	0.509	0.317
Subha Sawera Samma Ka sath	0.367	0.148	0.34	0.24
The Morning Show	0.117	0.389	0.075	0.253
Active Margin	1	1	1	

The above table shows the column profiles of the variable favorite morning show. For the column profiles, which has inertia of Jago Pakistan Jago is the highest .317, represents 48.9% of the total inertia, while the second highest The Morning Show has inertia of is 0.253, represents 27.1%. From Table 4 it can be seen that Jago Pakistan Jago has contribute 48.90% to this axis and it has 50% of the total inertia of the colume profile. That is, nearly all of the variation in the values of the favourate mornig show category is due to the significant difference among Jago Pakistan Jao, Good Mornnig Pakistan, The Morning Show and Subha Sawera Samma Ka Sath.

Table 5: Correspondence analysis model summary of the favorite morning shows and branded cloths

Model Summary					
Dimension	Cronbach's Alpha	Variance Accounted For			
		Total (Eigenvalue)	Inertia	% of Variance
1	.866	3.595	.599	59.915
2	.769	2.783	.464	46.384
Total		**6.378**	**1.063**	
Mean	.824a	3.189	.531	53.150
a. Mean Cronbach's Alpha is based on the mean Eigenvalue.				

Figure 1 shows that the variables between row and column are associated with each other on linearly manner, which means that the morning shows are making impact on the purchase of branded cloths in the market. It is showing a lineally relationship between the variables. Cronbach's alpha is a measure used to assess the reliability, or internal consistency, of a set of scale or test items. In other words, the reliability of any given measurement refers to the extent to which it is a consistent measure of a concept, and Cronbach's alpha is one way of measuring the strength of that consistency. It is computed by correlating the score for each scale item with the total score for each observation (usually individual survey respondents or test takers), and then comparing that to the variance for all individual item
In the above case Cronbach alpha is having 0.866 value for dimension 1, which is showing very high internal consistency. On the other hand, dimension 2 has the value 0.769 which is also a very high consistency within the variables.

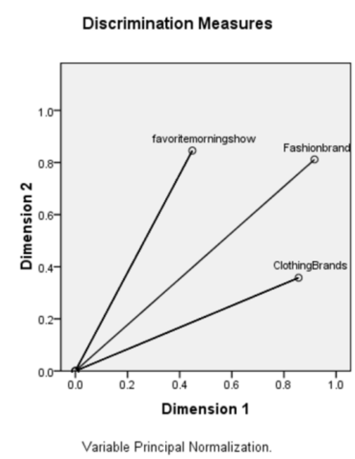

Fig. 2 Two discriminate measure plot of the association between the morning shows and brands

Above is the graphical presentation of row and columns variables using correspondence analysis, which defines the association among the variables of row and column profile. The coordinates of row and column contribution shows the d

5. Conclusion

Multiple correspondence analysis can be performed on the contingency table to visualize the association among variables. Correspondence analysis is a technique that allows the user to graphically display row and column categories and provide a visual inspection of their 'correspondences', or associations, at a category level. Therefore, when one concludes from a chi-squared test of independence that an association exists between the categorical variables correspondence analysis provides an intuitive graphical display of this association. Instead, the term simple refers to the fact that it is the simplest of contingency tables (consisting of only two cross-classified categorical variables the row variable and the column variable) that is being analyzed.

Such a figure represents an important component of the output generated from multiple correspondence analysis of the data in Table 1. It reveals that, in general, the view of morning shows has an impact of brands image and is associate with the purchase of branded cloth due to the celebrity endorsement on these products. Furthermore, from the plot, it is found that the first dimension visualizes 52.4% of the association (as described by the chi-squared statistic) between the row and column categories, whereas the second dimension visualizes 4.0%. Chi square tests were conducted and it was found that the associations among variables are significant. Another advantage of this method is that it can visualize the importance of a profile on a particular axis. This then means the contribution a particular profile makes with a component can be calculated and visualized. Therefore, the correspondence plot provides an excellent visual summary of the association between the variables.

References

[1] Agresti, A., 1982 : Ordinal Data, In Encyclopedia of Statistical Sciences, Vol 6, 511-516.

[2] Agresti, A., 1983 : A survey of strategies for modeling cross-classifications having ordinal variables, Journal of the American Statistical Association, 78, 184-198.

[3] Beh, E. J., 1996a : Correspondence analysis of two-way singly-ordered contingency tables, Department of Applied Statistics, University of Wollongong, Preprint No. 1/96.

[4] Beh, E. J., 1996b : Correspondence analysis of two-way doubly-ordered contingency tables, Department of Applied Statistics, University of Wollongong, Preprint No. 2/96.

[5] de Leeuw, J., 1993 : Some generalisations of correspondence analysis, In Multivariate Analysis: Future Directions 2, (ed Cuadras, C. M . & Rao, C. R.), pg 359-375.

[6] de Leeuw, J. & van der Heijden, 1988 : Correspondence analysis of incomplete contingency tables, Psychometrika, 53, 223-233.

[7] Gifi, A., 1990 : Non-linear Multivariate Analysis, John Wiley & Sons, Chichester

[i] Beh Eric.J Correspondence analysis using orthogonal polynomial (1998)

[ii] Beh Eric.J Correspondence analysis using orthogonal polynomial (1998)