Quantitative Methods for SOCIAL SCIENCES

Edited by: Vinayak Nikam • Abimanyu Jhajhria • Suresh Pal

This reference book is designed keeping in mind the need for the application of advanced quantitative methods in social science research to enhance its accuracy. The chapters are written in such a way that social scientists can easily grasp the methods including their theoretical and practical aspects using statistical software. The book provides comprehensive coverage of multivariate techniques, forecasting methods, structural equations, optimization models, quantitative methods for impact assessment, growth models and other important methods used in social science research.

ABOUT THE EDITORS

Vinayak Nikam is working as Scientist (Senior Scale) at ICAR-National Institute of Agricultural Economics and Policy Research (New Delhi). He did his PhD from ICAR-Indian Agricultural Research Institute (New Delhi) in agricultural extension. He also served ICAR-Central Soil Salinity Research Institute (Karnal) before joining NIAP in Nov 2015. He also holds faculty membership in the discipline of Agricultural Extension at ICAR-Indian Agriculture Research Institute (New Delhi). Currently, he is working on the performance and impact assessment of Agriculture Extension and Advisory Services as well as associated with technology impact assessment.

Abimanyu Jhajhria is a Scientist (Agricultural Economics) at ICAR-National Institute of Agricultural Economics and Policy Research (New Delhi). He has received his doctorate in Agricultural Economics from the ICAR-Indian Agricultural Research Institute (New Delhi). His research interest includes markets, trade and institutions. He is involved in the projects on market reforms, agricultural commodity value chains and outlook models for agricultural commodities.

Suresh Pal is Director of ICAR-National Institute of Agricultural Economics and Policy Research (New Delhi). He has a PhD in agricultural economics from ICAR-Indian Agricultural Research Institute (New Delhi). He has published extensively on different aspects of Indian agriculture and guided doctoral and masters students. He has received awards for his contributions like best journal article awards, Norman E Borlaug International Science Fellowship, Fellow of the Indian Society of Agricultural Economics (Mumbai), and Fellow of the National Academy of Agricultural Sciences.

ISBN:978-81-940080-2-6

ICAR-NATIONAL INSTITUTE OF AGRICULTURAL ECONOMICS AND POLICY RESEARCH (NIAP)
D.P.S.Marg, Pusa, New Delhi - 110012. Phone: +91-11- 25847628, +91-11- 25848731
Fax: +91-11-25842684 . E-mail:director.niap@icar.gov.in; Website: www.http://www.ncap.res.in

Quantitative Methods for Social Sciences Vinayak Nikam · Abimanyu Jhajhria · Suresh Pal

Vinayak Nikam • Abimanyu Jhajhria • Suresh Pal

Quantitative Methods for SOCIAL SCIENCES

ICAR-National Institute of Agricultural Economics and Policy Research

Reference book on

Quantitative Methods for SOCIAL SCIENCES

Edited by

Vinayak Nikam
Abimanyu Jhajhria
Suresh Pal

भाकृअनुप
ICAR

ICAR-National Institute of Agricultural Economics and Policy Research
New Delhi

Chapter 4

MULTIDIMENTIONAL SCALING

Ramasubramanian V.

INTRODUCTION

Multi-dimensional scaling (MDS) is basically a data visualization method that helps the analyst to uncover the hidden structure in multidimensional data. While a number of MDS methods exist with a plethora of different algorithms to arrive at MDS plots, each has been designed to arrive at an optimal low-dimensional configuration (usually in two or three dimensions) for the data under consideration. Some early applications for which MDS has been used were to generate a perceptual map given the (perceived) preferences/ similarities/ dissimilarities of objects subjected to pairwise comparison by individuals and also to reconstruct an original map given a two-way table of (objective) proximities/ distances of places (say, cities), instrumental in paving a way for a greater understanding of the power of MDS. Needless to say, over time, MDS has been extensively used in various application areas such as psychology, marketing, molecular biology and bioinformatics, agriculture etc. In these applications, in some cases, MDS has been used to identify key dimensions underlying respondents' evaluations of objects. From another standpoint, MDS has been used for identifying clusters of points, with points within a particular cluster viewed as being 'closer' to the other points in that cluster than to points present in other clusters. The theoretical concept of MDS, its various methods (as it is not a single procedure), elaboration of the analytical methods involved are discussed along with applications of the different types of MDS employed illustrated with the help of examples. In this chapter, these MDS procedures have been demonstrated on various datasets by analysing them in R software and hence the R codes and corresponding output with interpretation are also given (along with datasets used).

A detailed discussion regarding various MDS techniques can be found in many books (Kruskal and Wish, 1978; Chatfield and Collins, 1980; Coxon, 1982; Hair *et al.*, 1995; Cox and Cox, 2001; Borg and Groenen, 2005; Izenman, 2008). de Leeuw and Mair (2009) have given a good account on MDS discussing the various versions of MDS in a lucid manner and also R software MDS package named SMACOF i.e. *Scaling by MAjorising a COmplicated Function* in which all the known MDS procedures are embedded. Practical applications of MDS has been done by many researchers. Vishwanath and Chen (2006) have examined empirically the composition of technology clusters of several technology concepts and the differences in these clusters formed by adopters and non-adopters using the Galileo system of MDS and the associational diffusion framework. Ramasubramanian *et al.* (2014), while envisioning future technological needs for plant breeding and genetics subdomain of Indian agriculture, employed MDS using information obtained from experts for identifying key agricultural dimensions emerging out from the factors responsible relating to prioritizing new crop varieties and found that the factors seem to cluster together into two genetic and environmental groups.

TYPES AND METHODS OF MDS

The way MDS is done has evolved over years. Usually 'objects' (they could be tangible or intangible) are compared pairwise ('all pairs design') by different 'subjects' (they could be persons rating the objects or these could be repeated observations of same objects). Again the situations under which the data are collected could differ such as experimental conditions, subjects, stimuli etc. Accordingly, the two broad types of MDS are given as follows:

(i) One way or multi-way (i.e. K) MDS: In case of K (>1) way MDS, each pair of objects has K dissimilarity measures one each from different 'replications' like repeated measures, multiple raters etc.

(ii) One mode or multi-mode MDS: Here the K (>1) modes have their dissimilarities qualitatively different like experimental conditions, subjects, stimuli etc.

Also note that each of the above versions has metric and non-metric versions. An example of the K-way MDS is the INDSCAL i.e. INdividual Differences SCALing wherein K separate n x n symmetric dissimilarity matrices are there from the K judges for n objects. For more details, see Borg and Groenen (2005).

Classical MDS

Depending on the way the data are collected, various proximity (read similarity)/ dissimilarity measures can be taken into consideration. In this section, let us discuss the classical MDS under metric scaling set up. The conventional form of performing MDS is nothing but classical MDS wherein the dissimilarities (always it is better to convert similarities into dissimilarities for convenient representation purposes) can be taken as (say, Euclidean) distances without any additional transformation. Almost any standard book or chapter that attempts to give an exposition on MDS usually starts with a certain 'golden oldie' problem which is discussed subsequently. It is obvious that given a map, with cities marked in that, a matrix of pairwise distances (in MDS parlance, 'distances' are likened to 'dissimilarities') between these cities can always be formed. But suppose, the problem is looked the other way round, i.e. only the pairwise distances between these cities are available and the map has to be constructed. A 'perceptual map' solution is possible by the classical MDS approach under metric scaling (as the distances here are actual measurements and not on any gradation on, say, a five point scale, in which case it becomes non-metric; the latter case is encountered when there is a perception/ preference/ dis/similarity matrix) and sometimes also known as 'principal coordinates' analysis. For this, consider the following matrix of distances (in kilometres) among 21 European cities (given in R software itself) Table 1.

Table 1: Distances of European cities in kilometres

↓ City →	1.Athens	2	3	4	5	6	7	8	9	10
2.Barcelona	3313									
3.Brussels	2963	1318								
4.Calais	3175	1326	204							
5.Cherbourg	3339	1294	583	460						
6.Cologne	2762	1498	206	409	785					
7.Copenhagen	3276	2218	966	1136	1545	760				
8.Geneva	2610	803	677	747	853	1662	1418			
9.Gibraltar	4485	1172	2256	2224	2047	2436	3196	1975		
10.Hamburg	2977	2018	597	714	1115	460	460	1118	2897	
11.Hook of Holland	3030	1490	172	330	731	269	269	895	2428	550
12.Lisbon	4532	1305	2084	2052	1827	2290	2971	1936	676	2671
13.Lyons	2753	645	690	739	789	714	1458	158	1817	1159
14.Madrid	3949	636	1558	1550	1347	1764	2498	1439	698	2198
15.Marseilles	2865	521	1011	1059	1101	1035	1778	425	1693	1479
16.Milan	2282	1014	925	1077	1209	911	1537	328	2185	1238
17.Munich	2179	1365	747	977	1160	583	1104	591	2565	805
18.Paris	3000	1033	285	280	340	465	1176	513	1971	877
19.Rome	817	1460	1511	1662	1794	1497	2050	995	2631	1751
20.Stockholm	3927	2868	1616	1786	2196	1403	650	2068	3886	949
21.Vienna	1991	1802	1175	1381	1588	937	1455	1019	2974	1155
↓ City →	11	12	13	14	15	16	17	18	19	20
12.Lisbon	2280									
13.Lyons	863	1178								
14.Madrid	1730	668	1281							
15.Marseilles	1183	1762	320	1157						
16.Milan	1098	2250	328	1724	618					
17.Munich	851	2507	724	2010	1109	331				
18.Paris	457	1799	471	1273	792	856	821			
19.Rome	1683	2700	1048	2007	1011	586	946	1476		
20.Stockholm	1500	3231	2108	3188	2428	2187	1754	1827	2707	
21.Vienna	1205	2937	1157	2409	1363	898	428	1249	1209	2105

Using the above data with the following R code, the output obtained is given in Fig 1.

 # In the following set of R codes, the statements after '#' are comments

loc <- cmdscale(eurodist,k=2,) *#'classical MDS' with the dataset in 'eurodist' (distances #between European cities within datasets package of R software and 'loc' is name given from the user's side for storing the result output in R*

x <- loc[, 1] *# the first coordinate values for the 21 cities* y <- -loc[, 2] # *the second coordinate values for the 21 cities; a 'reflection' is done by adding a negative sign so that North direction is at the top in the map*

plot(x, y, type = "n", xlab = "", ylab = "", asp = 1, axes = FALSE,

main = "cmdscale(eurodist)") ## note asp = 1, to ensure aspect ratio of y/x is 1 so that Euclidean distances are represented correctly

text(x, y, rownames(loc), cex = 1.0)

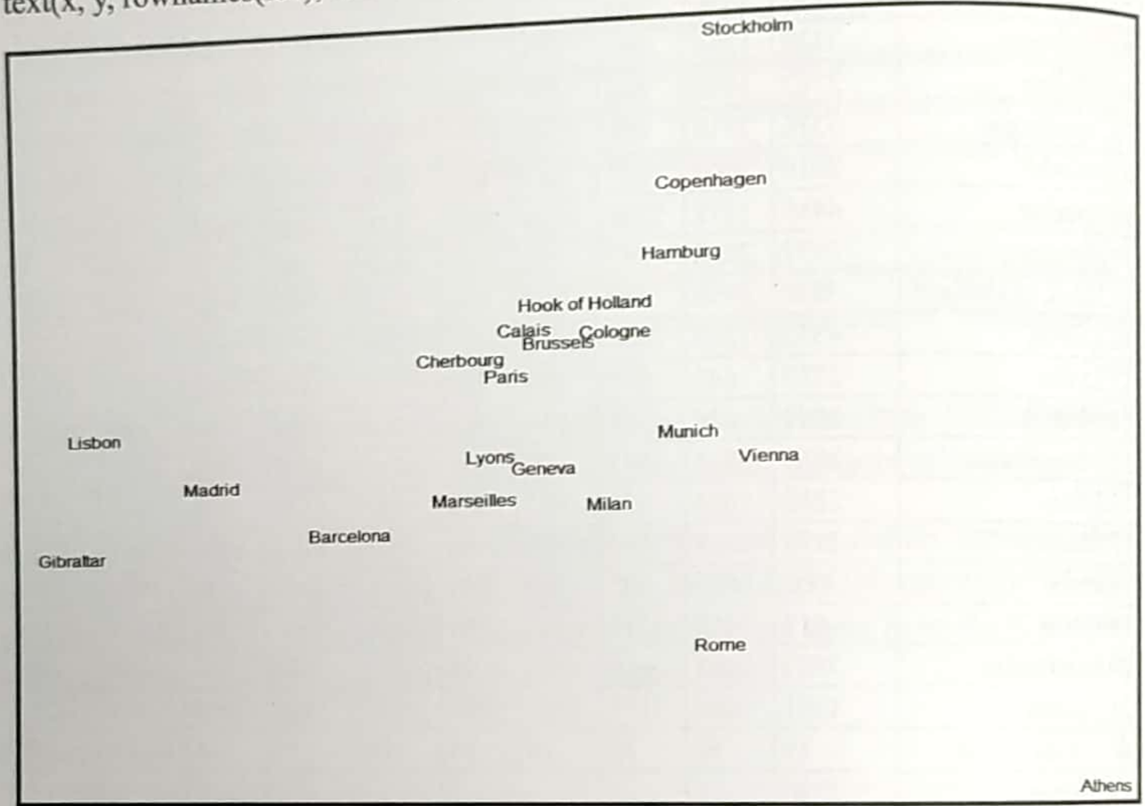

Fig 1: The MDS map constructed using the distances among 21 European cities

The MDS map in Fig 1 can be compared with the actual map given in Fig 2. It can be seen that, by and large, the map has been reconstructed. However, one should also take into account the reflection and rotation of the map. The line distances between cities are represented in the MDS plot while the map involves the curvature represented in a plane. Thus, some distortions can be noticed, for example, the positions of the two cities Madrid and Lisbon in the MDS plot as compared to the original map. Such errors would arise because the distances between cities would in reality need to be represented in more than two dimensions. Nevertheless, essentially what MDS has tried to generate is to find the coordinates '(x,y}' so that the same can be represented in a map. It is also noted here that Anonymous (2007) has attempted to construct the same using built-in Solver Add-in utility of MS Excel by first principle approach with the steps involved explained in a lucid manner. For this, a starting configuration for the 'n' objects (n=21 'cities' in the above discussion) in the two dimensions i.e. coordinates (x, y) were arbitrarily selected for each object. The Euclidean distances (d_{ij}'s) between the objects in the arbitrarily constructed plot are then computed. Let δ_{ij}'s be the distances between i and j indicated in the original diagonal matrix (input data).

Any theory on MDS, say, as given in Manly and Alberto (2017), proceeds to have a regression of d_{ij} on δ_{ij} whose form f(.) can be linear, polynomial, or monotonic. The fitted distances are called 'disparities' $\hat{d}_{ij}$. (Note here that no such transformation is done in the above city distances illustration, as the measurement is metric). These disparities $\hat{d}_{ij}$ are nothing but the data distances δ_{ij} scaled to match the configuration distances d_{ij} as closely as possible. The goodness of fit between the configuration distances and the disparities is measured by a suitable statistic. One common measure is the Kruskal's stress formula 1 given by

$$\text{Stress } 1 = \frac{\sum\left(d_{ij} - \hat{d}_{ij}\right)^2}{\sum \hat{d}_{ij}^2}$$

This statistic is a measure of the extent to which the spatial configuration of points has to be stressed in order to obtain the data distances δ_{ij}, hence the name 'stress'. Note that the less the stress, the more the good is the fit. It is noted here that the Σ (summation) both in the numerator and the denominator extends over all the 210 ($= {}_{21}C_2$) pairwise distances between the 21 cities.

Thus from a mathematical point of view, what MDS does is to find a set of vectors in p-dimensional space (p=2 in the above discussion) such that the matrix of Euclidean distances among them corresponds as closely as possible to some function (identity function in the above discussion) of the input matrix according to a criterion function like Stress 1.

Fig 2: The actual map of the 21 European cities considered (Source: Google)

Now, let us consider another example. Consider the case of the matrix of distances (Table 2) among cities of United States (given in R software itself) and MDS produces a map as shown in Fig 3 which can be compared with the actual map given in Fig 4.

Table 2: Distance between places of United States (in mile)

	Boston	NY	DC	Miami	Chicago	Seattle	SF	LA	Denver
Boston	0	206	429	1504	963	2976	3095	2979	1949
NY	206	0	233	1308	802	2815	2934	2786	1771
DC	429	233	0	1075	671	2684	2799	2631	1616
Miami	1504	1308	1075	0	1329	3273	3053	2687	2037
Chicago	963	802	671	1329	0	2013	2142	2054	996
Seattle	2976	2815	2684	3273	2013	0	808	1131	1307
SF	3095	2934	2799	3053	2142	808	0	379	1235
LA	2979	2786	2631	2687	2054	1131	379	0	1059
Denver	1949	1771	1616	2037	996	1307	1235	1059	0

The necessary R codes are given below:

```
df <- read.csv(file.choose(),header = TRUE)

row.names(df) <- df[, 1]

df <- df[, -1]

fit <- cmdscale(df, eig = TRUE, k = 2)

x <- fit$points[, 1]

y <- fit$points[, 2]

plot(x, y, pch = 19, xlim = range(x) + c(0, 600))

city.names <- c("BOSTON","NY","DC","MIAMI","CHICAGO","SEATTLE","SF
","LA", "DENVER")

text(x, y, pos = 4, labels = city.names)

x <- 0 - x

y <- 0 - y

plot(x, y, pch = 19, xlim = range(x) + c(0, 600))

text(x, y, pos = 4, labels = city.names)
```

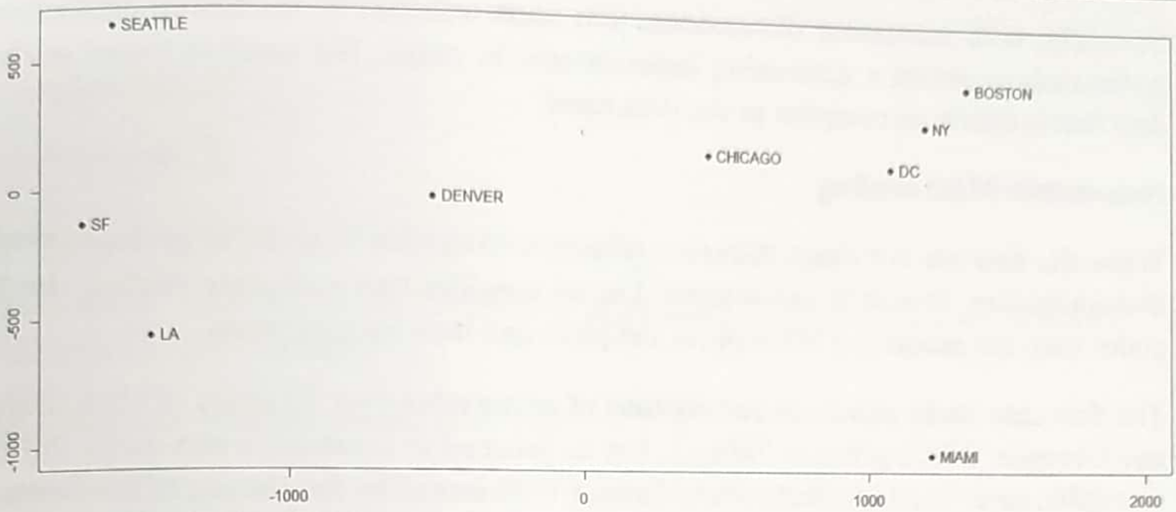

Fig 3: The MDS map of the nine places of United States considered

Normally, MDS is used to provide a visual representation of a complex set of relationships that can be scanned at a glance. Since maps on paper are two-dimensional objects, this translates technically to finding an optimal configuration of points in 2-dimensional space. However, the best possible configuration in two dimensions may be a very poor, highly distorted, representation of your data. When this happens, you can increase the number of dimensions, if required. There are two difficulties with increasing the number of dimensions. Firstly, even three dimensions are difficult to display on paper and are significantly more difficult to comprehend. Four or more dimensions' render MDS virtually useless as a method of making complex data more accessible to the human mind.

Fig 4: The actual map of the nine places of United States considered (Source: Google)

Secondly, with increasing dimensions, you must estimate an increasing number of parameters to obtain a decreasing improvement in stress. The result is model of the data that is nearly as complex as the data itself.

Non-metric MDS scaling

When the data are not exact distances (dissimilarities) but in terms of perceptions or dis/similarities, then it is non-metric. Let us consider two such case studies, the R codes used for producing MDS plots, the plots and their interpretation.

The first case study relates to correlations of crime rates over 50 states of USA (Borg and Groenen, 2005) given in Table 3. Let us proceed to construct a non-metric MDS using this data. The U.S. Statistical Abstract 1970 issued by the Bureau of the Census provides statistics on the rate of different crimes in the 50 U.S. states. One question that can be asked about these data is to what extent one can predict a high crime rate of murder, say, by knowing that the crime rate of burglary is high. A partial answer to this question is provided by computing the correlations of the crime rates over the 50 U.S. states (Table 3). But even in such a fairly small correlation matrix, it is not easy to understand the structure of these coefficients. This task is made much simpler by representing the correlations in the form of a "picture" of a two-dimensional MDS representation given in Fig 5 where each crime is shown as a point. The points are arranged in such a way that their distances correspond to the correlations. That is, two points are close together (such as murder and assault) if their corresponding crime rates are highly correlated. Conversely, two points are far apart if their crime rates are not correlated that highly (such as assault and larceny).

Table 3: Correlations of crime rates over 50 states of USA

	Murder	Rape	Robbery	Assault	Burglary	Larseny	Autotheft
Murder	1	0.52	0.34	0.81	0.28	0.06	0.11
Rape	0.52	1	0.55	0.7	0.68	0.6	0.44
Robbery	0.34	0.55	1	0.56	0.62	0.44	0.62
Assault	0.81	0.7	0.56	1	0.52	0.32	0.33
Burglary	0.28	0.68	0.62	0.52	1	0.8	0.7
Larseny	0.06	0.6	0.44	0.32	0.8	1	0.55
Autotheft	0.11	0.44	0.62	0.33	0.7	0.55	1

The necessary R codes are given below:

```
df <- read.csv(file.choose(),header = TRUE)

row.names(df) <- df[, 1]

df <- 1-df[, -1]
```

```
fit <- cmdscale(df,add=TRUE)

x <- fit$points[, 1]

y <- fit$points[, 2]

plot(x, y,asp = 1, axes = FALSE)

crimes.names <- ("MURDER","RAPE","ROBBERY","ASSAULT","BURGLARY",
"LARSENY","AUTOTHEFT")

text(x, y, labels = crimes.names, cex=1.2)
```

Once the MDS plot is obtained, by visual inspection, the dimensions of the perceptual map can be labelled as given in Fig 5.

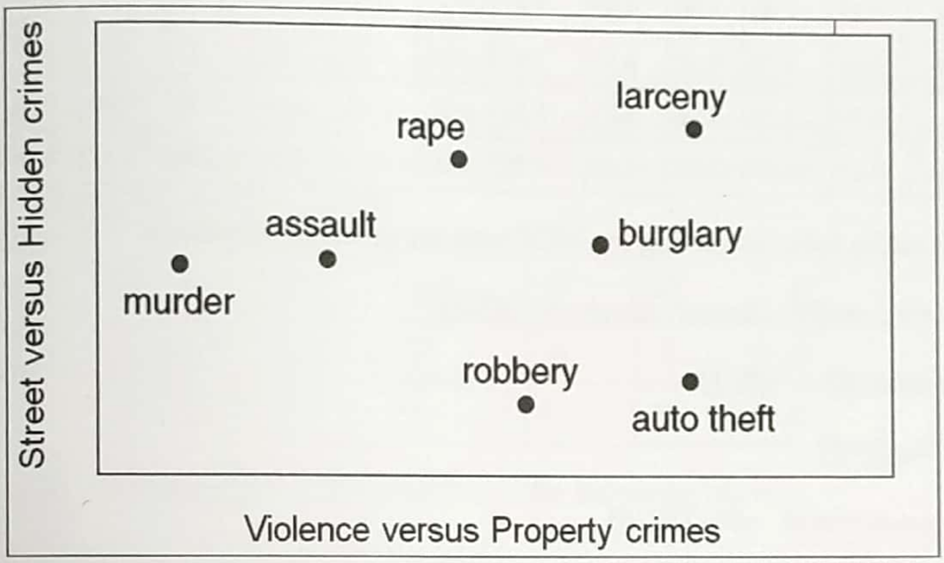

Fig 5: MDS map of crimes over 50 states of USA

The second case study relates to Similarity ratings for 12 nations (rated by individuals) given in Table 4. The data are from a pilot study on perceptions of nations (Kruskal and Wish, 1978). Each of the 18 students (in a psychological measurement course taught by Wish) participating in the study rated the degree of overall similarity between twelve nations on a scale ranging from 1 for "very different" to 9 for "very similar." There were no instructions concerning the characteristics on which these similarity judgments were to be made; this was information to discover rather than to impose. The first step of the data analysis was to compute the mean similarity rating for each of the 66 pairs (all combinations of the 12 nations i.e. ${}_{12}C_2$). Thus, for example, USSR and Yugoslavia were perceived to be more similar to each other (mean = 6.67) than any other pair of nations, while China-Brazil and USA-Congo were judged to be the most dissimilar pairs (mean = 2.39). The corresponding MDS plot is given in Fig 6) which also have the dimensions labeled by the researcher.

Table 4: Similarity ratings for 12 nations (rated by individuals)

	Brazil	Congo	Cuba	Egypt	France	India	Israel	Japan	China	USSR	USA	Yugosl.
Brazil	9.00	4.83	5.28	3.44	4.72	4.50	3.83	3.50	2.39	3.06	5.39	3.17
Congo	4.83	9.00	4.56	5.00	4.00	4.83	3.33	3.39	4.00	3.39	2.39	3.50
Cuba	5.28	4.56	9.00	5.17	4.11	4.00	3.61	2.94	5.50	5.44	3.17	5.11
Egypt	3.44	5.00	5.17	9.00	4.78	5.83	4.67	3.83	4.39	4.39	3.33	4.28
France	4.72	4.00	4.11	4.78	9.00	3.44	4.00	4.22	3.67	5.06	5.94	4.72
India	4.50	4.83	4.00	5.83	3.44	9.00	4.11	4.50	4.11	4.50	4.28	4.00
Israel	3.83	3.33	3.61	4.67	4.00	4.11	9.00	4.83	3.00	4.17	5.94	4.44
Japan	3.50	3.39	2.94	3.83	4.22	4.50	4.83	9.00	4.17	4.61	6.06	4.28
China	2.39	4.00	5.50	4.39	3.67	4.11	3.00	4.17	9.00	5.72	2.56	5.06
USSR	3.06	3.39	5.44	4.39	5.06	4.50	4.17	4.61	5.72	9.00	5.00	6.67
USA	5.39	2.39	3.17	3.33	5.94	4.28	5.94	6.06	2.56	5.00	9.00	3.56
Yugosl.	3.17	3.50	5.11	4.28	4.72	4.00	4.44	4.28	5.06	6.67	3.56	9.00

The R codes for constructing the MDS map are given subsequently.

```
df <- read.csv(file.choose(),header = TRUE)

row.names(df) <- df[, 1]

df <- 9-df[, -1]

fit <- cmdscale(df, add=TRUE)

x <- fit$points[, 1]

y <- fit$points[, 2]

plot(x, y,asp = 1, axes = FALSE)

country.names<- c("Brazil","Congo","Cuba","Egypt","France","India","Israel",
"Japan","China","USSR","USA","Yugoslavia")

text(x, y, labels = country.names, cex=1)
```

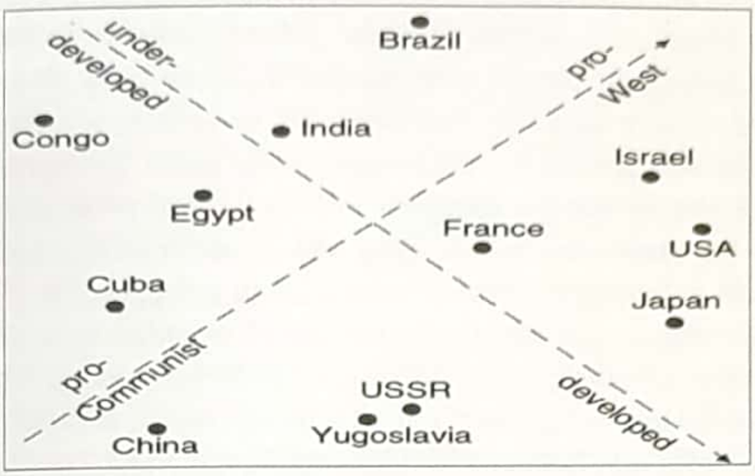

Fig 6: MDS map for similarity ratings for 12 nations (rated by individuals)

Unfolding

One more variant of MDS is the 'unfolding' method which requires rectangular MDS input matrices of order (n_1 x n_2) wherein n_1 judges rate n_2 objects. Suppose n_1 =2 judges (1 and 2) consider a set of n_2=5 objects (say research papers A, B, C, D, E) and individually rank them as follows:

	1st	2nd	3rd	4th	5th
Judge 1	B	C	A	E	D
Judge 2	A	B	C	E	D

The above observations can be arranged in the following manner:

D	E		C		1	B		2	A

In the following arrangement, it can be seen that the distances from judge 1 to the five 'objects' have the same ranking as his original ranking of the objects; similarly for judge 2.

1	B		C	A	E	D
2	A	B		C	E	D

For each judge the line can be folded at the judge's position and their original rankings are observed. Alternatively, looking at the situation in reverse, the judges' rankings when placed on a line can be 'unfolded' to obtain the "common" ordering of the objects.

Thus unfolding tries to model preferential choice by assuming that different individuals perceive the various objects of choice in the same way but differ with respect to what they consider an ideal combination of the objects' attributes. In unfolding, the data are preference scores such as rank-orders of different individuals for a set of choice objects. These data can be conceived as proximities between the elements of two sets, individuals and objects. In a way, unfolding can be seen as a special case of MDS where the within-sets proximities are missing. Individuals are represented as 'ideal' points in MDS plot so that the distances from each ideal point to the object points correspond to the preference scores. Borg and Groenen (2005) has introduced the basic notions of unfolding models by means of an example. They considered data wherein 42 individuals were asked to rank-order 15 breakfast items (A=toast pop-up; B=buttered toast; C=English muffin and margarine; D=jelly donut; E=cinnamon toast; F=blueberry muffin and margarine; G=hard rolls and butter; H=toast and marmalade; I=buttered toast and jelly; J=toast and margarine; K=cinnamon bun; L=Danish pastry; M=glazed donut; N=coffee cake; O=corn muffin and butter) from 1 (= most preferred) to 15 (= least preferred) which are given in Table 5. In Table 5, each row contains the ranking numbers assigned to breakfast items A, ..., O by individual i (=1 to 42).

Fig 7 presents an unfolding solution to the above data. The resulting configuration consists of 57 points, 42 for the individuals (shown as numbers) and 15 for the breakfast items (shown as their short names). Every individual is represented by an ideal point. The closer an object point lies to an ideal point, the more the object is preferred by the respective individual.

Table 5: Individual rank-order for breakfast items (Unfolding example)

	Breakfast items														
Individual↓	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**
1.	13	12	7	3	5	4	8	11	10	15	2	1	6	9	14
2.	15	11	6	3	10	5	14	8	9	12	7	1	4	2	13
3.	15	10	12	14	3	2	9	8	7	11	1	6	4	5	13
4.	6	14	11	3	7	8	12	10	9	15	4	1	2	5	13
5.	15	9	6	14	13	2	12	8	7	10	11	1	4	3	5
6.	9	11	14	4	7	6	15	10	8	12	5	2	3	1	13
7.	9	14	5	6	8	4	13	11	12	15	7	2	1	3	10
8.	15	10	12	6	9	2	13	8	7	11	3	1	5	4	14
9.	15	12	2	4	5	8	10	11	3	13	7	9	6	1	14
10.	15	13	10	7	6	4	9	12	11	14	5	2	8	1	3
11.	9	2	4	15	8	5	1	10	6	7	11	13	14	12	3
12.	11	1	2	15	12	3	4	8	7	14	10	9	13	5	6
13.	12	1	14	4	5	6	11	13	2	15	10	3	9	8	7
14.	13	11	14	5	4	12	10	8	7	15	3	2	6	1	9
15.	12	11	8	1	4	7	14	10	9	13	5	2	6	3	15
16.	15	12	4	14	5	3	11	9	7	13	6	8	1	2	10

	Breakfast items														
Individual↓	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**
17.	7	10	8	3	13	6	15	12	11	9	5	1	4	2	14
18.	7	12	6	4	10	1	15	9	8	13	5	3	14	2	11
19.	2	9	8	5	15	12	7	10	6	11	1	3	4	13	14
20.	10	11	15	6	9	4	14	2	13	12	8	1	3	7	5
21.	12	1	2	10	3	15	5	6	4	13	7	11	8	9	14
22.	14	12	10	1	11	5	15	8	7	13	2	6	4	3	9
23.	14	6	1	13	2	5	15	8	4	12	7	10	9	3	11
24.	10	11	9	15	5	6	12	1	3	13	8	2	14	4	7
25.	15	8	7	5	9	10	13	3	11	6	2	1	12	4	14
26.	15	13	8	5	10	7	14	12	11	6	4	1	3	2	9
27.	11	3	6	14	1	7	9	4	2	5	10	15	13	12	8
28.	6	15	3	11	8	2	13	9	10	14	5	7	12	1	4
29.	15	7	10	2	12	9	13	8	5	6	11	1	3	4	14
30.	15	10	7	2	9	6	14	12	8	11	5	3	1	4	13
31.	11	4	9	10	15	8	6	5	1	13	14	2	12	3	7
32.	9	3	10	13	14	11	1	2	4	5	15	6	7	8	12
33.	15	8	1	11	10	2	4	13	14	9	6	5	12	3	7
34.	15	8	3	11	10	2	4	13	14	9	6	5	12	1	7
35.	15	6	10	14	12	8	2	4	3	5	11	1	13	7	9
36.	12	2	13	11	9	15	3	1	4	5	6	8	10	7	14
37.	5	1	6	11	12	10	7	4	3	2	13	9	8	14	15
38.	15	11	7	13	4	6	9	14	8	12	1	10	3	2	5
39.	6	1	12	5	15	9	2	7	11	3	8	10	4	14	13
40.	14	1	5	15	4	6	3	8	9	2	12	11	13	10	7
41.	10	3	2	14	9	1	8	12	13	4	11	5	15	6	7
42.	13	3	1	14	4	10	5	15	6	2	11	7	12	8	9

The R codes for constructing the MDS map are given subsequently. It is noted here that the SMACOF package of R has to be installed and loaded before proceeding for running these codes.

```
res <- unfolding(breakfast)

res

plot(res)
```

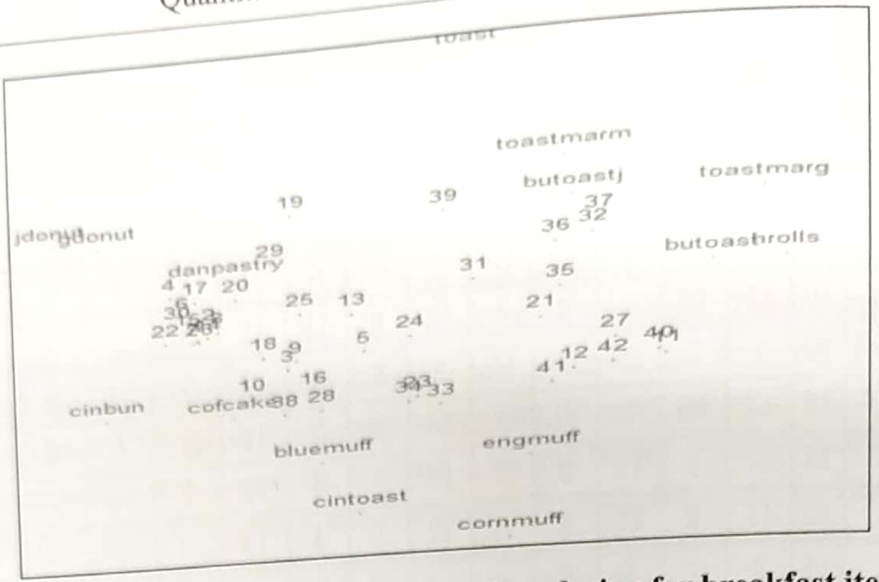

Fig 7: MDS unfolding of individual rank-ordering for breakfast items

Assume that Fig 7 was printed on a thin handkerchief. If this handkerchief is picked up with two fingers at the point representing individual i, say, y_i on the MDS plot and then pulled through the other hand, we have folded it: point y_i is on top, and the farther down the object points, the less preferred the objects they represent. The order of the points in the vertical direction corresponds (if we folded a perfect representation) to how individual i ordered these objects in terms of preference. Picking up the handkerchief in this way at any individual's ideal point yields that individual's empirical rank-order. The MDS process, then, is the inverse of the folding, that is, the unfolding of the given rank-orders into the distances. For example, individual 4 prefers L=Danish pastry (written 'danpastry' in the plot) the most, because the object points of these breakfast items are closest to this individual's ideal point, followed by M=glazed donut (gdonut) and D=jelly donut (jdonut), and then K=cinnamon bun ("cinbun") and N=coffee cake ("cofcake"). It is noted that for individual 4, J=toast and margarine ("toastmarg") is least preferred Somewhat less preferred is the coffee and cake breakfast (N), whereas A, B, F, I, H, C, G and O are more or less equally disliked. Now, one can confirm this plot ranking by referring to item ranking in the Table 5 against individual 4. In a similar manner, the ranking preferences of breakfast items for other individuals are also unfolded.

In this way, MDS can be used to understand and also unravel the information in any underlying data. While no claim is made to have discussed all the methods under MDS, the essential methods have been truthfully covered.

REFERENCES

Anonymous-Solver-MSExcel (2007), Constructing Perceptual Maps with the Aid of SOLVER in Office. http://cw.routledge.com/textbooks/9780415458160/instructorresources/MDS%20with%20spreadsheet.docx, accessed on 01 September, 2019.

Borg, I. and P. J. F. Groenen (2005), Modern Multidimensional Scaling: Theory and Applications, Second edition, *Springer-Verlag*, New York.

Chatfield, C. and A. J. Collins (1980), Introduction to Multivariate Analysis, Chapman and Hall, London.

Cox, T.F. and M. A. A. Cox (2001), Multidimensional Scaling, Second edition. Chapman & Hall/CRC, Boca Raton.

Coxon, A. P. M. (1982), The User's guide to Multidimensional Scaling, Heinemann Educational Books, Great Britain.

de Leeuw, J. and P. Mair (2009), Multidimensional scaling using majorization: SMACOF in R, *Journal of Statistical Software*, 31(3): 1-30.

Hair, J. F., R. E. Anderson, R.L. Tatha, and W. C. Black (1995), Multivariate Data Analysis, 4th Edition, Prentice Hall, New Jersey.

Izenman, A. J. (2008), Modern Multivariate Statistical Techniques: Regression, Classification and Manifold Learning. Springer, New York.

Kruskal, J. B. and M. Wish (1978), Multidimensional Scaling, Series: Quantitative applications in the Social Sciences, Sage University Press, California.

Manly, B. F. J. and J. A. N. Alberto (2017), Multivariate statistical methods: A Primer, Fourth Edition, CRC Press, Taylor and Francis Group, Boca Raton, Florida.

Ramasubramanian, V., A. Kumar, K. V. Prabhu, V. K. Bhatia and P. Ramasundaram (2014), Forecasting technological needs and prioritizing factors in agriculture from plant breeding and genetics domain perspective: A review, *Indian Journal of Agricultural Sciences*, 84 (3): 311-316.

Vishwanath, A. and H. Chen (2006), Technology clusters: using multidimensional scaling to evaluate and structure technology clusters, *Journal of the American Society for Information Science and Technology*, 57(11): 1451-1460