"Data\_Chocolate\_allinterviews.xlsx" sheet = "AttributeRatingsStacked"

'data.frame': 500 obs. of 15 variables: Person , Product ,crunchy , creamy , sweet,chocolaty, healthful, calorie , rich , addiction , accessible , handy , wrapping , image ,commercial

Using Eucledean Distance

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **\*** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
| 2 | 3.06 |  |  |  |  |  |  |  |  |
| 3 | 2.02 | 2.87 |  |  |  |  |  |  |  |
| 4 | 2.96 | 2.31 | 2.59 |  |  |  |  |  |  |
| 5 | 4.42 | 2.62 | 3.05 | 3.03 |  |  |  |  |  |
| 6 | 1.57 | 3.11 | 0.91 | 2.74 | 3.75 |  |  |  |  |
| 7 | 1.95 | 2.42 | 1.80 | 1.37 | 3.35 | 1.74 |  |  |  |
| 8 | 4.09 | 1.92 | 3.06 | 2.55 | 1.63 | 3.61 | 2.88 |  |  |
| 9 | 2.83 | 2.58 | 1.42 | 2.25 | 2.50 | 1.89 | 1.79 | 2.10 |  |
| 10 | 2.33 | 2.59 | 1.22 | 1.78 | 2.90 | 1.43 | 1.19 | 2.61 | 0.96 |

isoMDS command

Name: Kruskal's Non-metric Multidimensional Scaling

Description : One form of non-metric multidimensional scaling

Why did we use Non-metric Scaling 🡪 Non Metric data: Data collected from binary scales, nominal scales and ordinal scales are jointly termed as non metric data, that is, they do not possess a meter with which distance between scale values can be measured.

This technique is resorted to when the original data are of categorical or similar type that have been observed on a scale where only ranking is important and not actual differences. (Ordinal Scaling)

Non-metric multidimensional scaling (NMDS) is an indirect gradient analysis approach which produces an ordination based on a distance or dissimilarity matrix. Unlike methods which attempt to maximise the variance or correspondence between objects in an ordination, NMDS attempts to represent, as closely as possible, the pairwise dissimilarity between objects in a low-dimensional space. Any dissimilarity coefficient or distance measure may be used to build the distance matrix used as input.

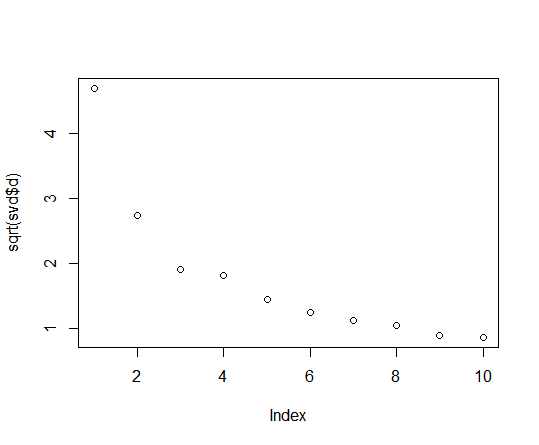
NMDS is a rank-based approach. This means that the original distance data is substituted with ranks. Thus, rather than object A being 2.1 units distant from object B and 4.4 units distant from object C, object C is the "first" most distant from object A while object C is the "second" most distant. While information about the magnitude of distances is lost, rank-based methods are generally more robust to data which do not have an identifiable distribution.

NMDS is a robust technique. It can:

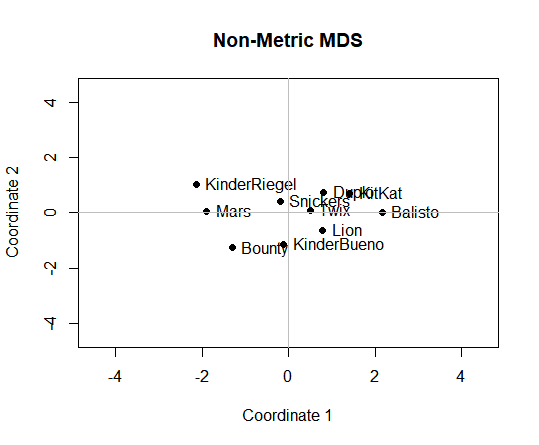
tolerate missing pairwise distances

be applied to a (dis)similarity matrix built with any (dis)similarity measure and

use quantitative, semi-quantitative, qualitative, or mixed variables



<https://mb3is.megx.net/gustame/dissimilarity-based-methods/nmds>



Points represent objects. Objects that are more similar to one another are ordinated closer together. The axes are arbitrary as is the orientation of the plot. Stress values should always accompany an NMDS ordination.

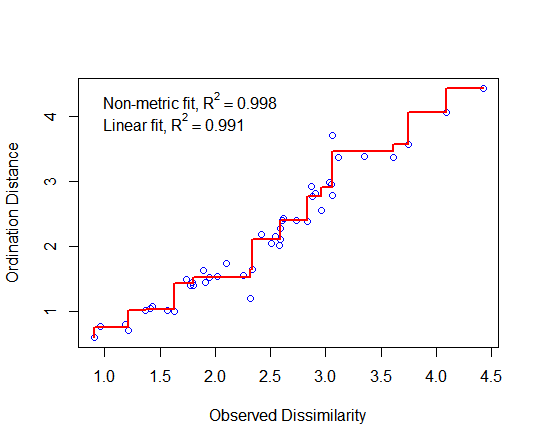
As noted above, the standard NMDS procedure focuses on accurately representing the distances in a distance matrix in an ordination. It does not attempt to create a visualisation which, for example, maximises the separation between points. Minimising overlap between points in an ordination is helpful in interpretation, thus a principal components analysis may be applied to NMDS axis scores. This effectively rotates an NMDS solution to ensure the first NMDS axis lies along the direction of maximum scatter. Several implementations of the NMDS routine do this automatically, however, performing this manually is not difficult and can be of great help.

**Reading NMDS plots is quite straightforward:** objects that are ordinated closer to one another are likely to be more similar than those further apart. However, the scale of the axes is arbitrary as is the orientation of the plot. Solutions with higher stress values (usually above 0.20) should be interpreted with caution and those with stress above 0.30 are highly suspect.Tight clusters of points that are well-separated from other clusters may indicate sub-populations in the data. Be aware that if a cluster of points is very dissimilar from other points or clusters, the arrangement of points in that cluster may not be very meaningful. The stress of the solution would be minimally affected by rearranging points in a tight cluster. Re-running an NMDS with only those objects in a given cluster may reveal more informative patterns.

NMDS is suited to indirect gradient analysis. If the patterns in an ordination corroborate existing knowledge or a hypothesis, there may be grounds for a direct gradient analysis, a hypothesis test, or the design of a new sampling campaign targeting that variation. Do not fall into the trap of data dredging, however.

REALLY GOOD ONE

<https://jonlefcheck.net/2012/10/24/nmds-tutorial-in-r/>



Large scatter around the line suggests that original dissimilarities are not well preserved in the reduced number of dimensions. Looks pretty good in this case. (According to the example at the website)

fit$stress

[1] 3.996328

High value of stress indicates

Choosing the appropriate number of dimensions is an interactive process. After an initial ordination, examine the stress values generated by the algorithm. As a rule of thumb, an NMDS ordination with a stress value around or above 0.2 is deemed suspect and a stress value approaching 0.3 indicates that the ordination is arbitrary. Stress values equal to or below 0.1 are considered fair, while values equal to or below 0.05 indicate good fit. Allowing the algorithm to ordinate in more dimensions can reduce the stress; however, allowing more than 3 dimensions quickly makes interpretation more challenging.

plot(fit)

