

The challenge with categorical variables is to find a suitable way to represent distances between variable categories and individuals in the factorial space. To overcome this problem, you can look for a non-linear transformation of each variable--whether it be nominal, ordinal, polynomial, or numerical--with optimal scaling. This is well explained in *Gifi Methods for Optimal Scaling in R: The Package homals*, and an implementation is available in the corresponding R package *homals*.

Instead using PCA we can use Multiple Correspondence Analysis.

In statistics, multiple correspondence analysis (MCA) is a data analysis technique for nominal categorical data, used to detect and represent underlying structures in a data set. It does this by representing data as points in a low-dimensional Euclidean space. The procedure thus appears to be the counterpart of principal component analysis for categorical data. MCA can be viewed as an extension of simple correspondence analysis (CA) in that it is applicable to a large set of categorical variables.