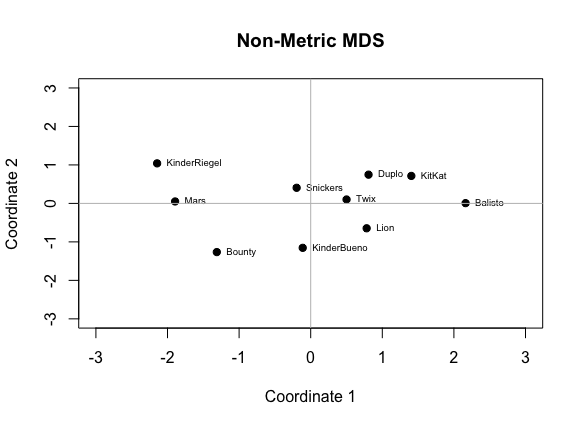
Special Work Performance 2

Giorgi Modebadze – Radoslav Evtimov – Ecenaz Bal

The purpose of this report is to analyze survey data of 50 different persons about 10 different chocolate brands in the German market. Each different brand is assessed with 13 different attributes evaluated on a scale 1 to 5(1-lowest,5-highest) We used Kruskal's Non-metric Multidimensional Scaling and Principal Component Analysis to make the data more interpretable.

SWP2a:

For the purposes of Non- Metric MDS, the missing values were replaced with the mean value for the quality of the given product. The reason for using non-metric scaling is that the original data is categorical where only ranking matters and not actual differences.To find how the different brands perceived, the Euclidean distance is used. Based on the result of Euclidean distance a two-dimentional perceptual map is created.



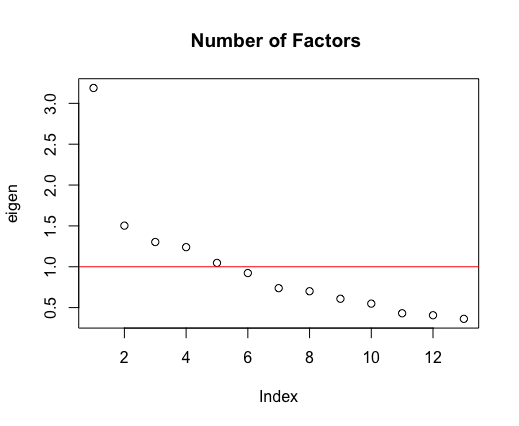
As it can be seen on the map, Balisto and KinderRiegel are the chocolate brands those have a lot more difference comparing to other brands. These brands are located far from each other which only says they are not similar.

However, it is not possible to interpret what attributes caused this dissimilarity unless doing Property Fitting. That is why, an another method is needed.

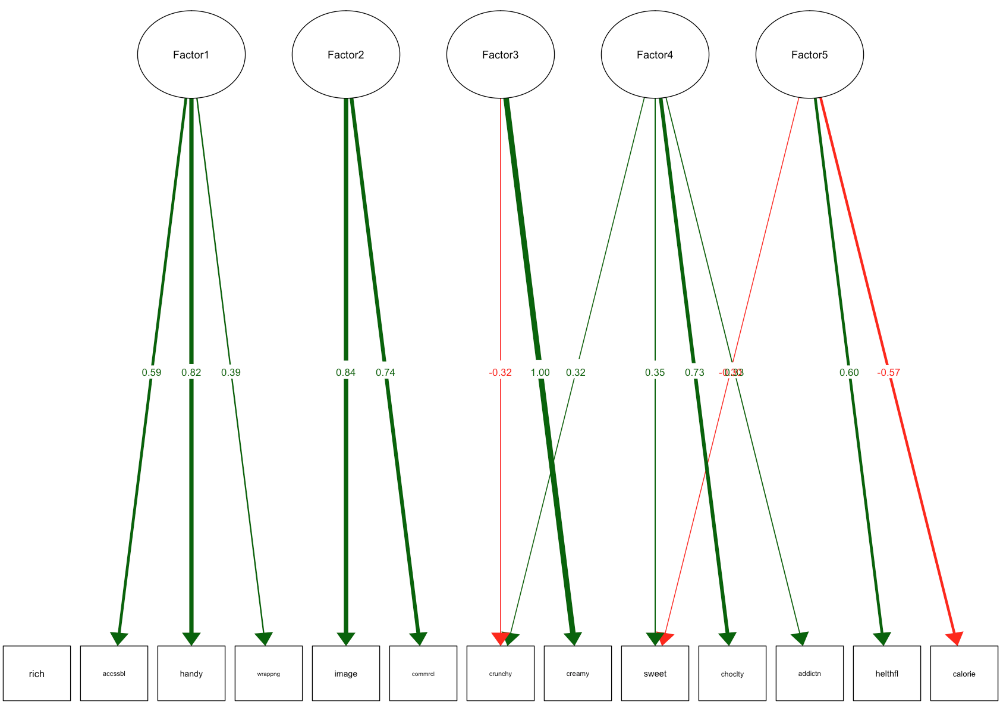
Running a regression of every attribute

To be able to interpret the position of the brand on the perceptual map, it is needed to know how the attributes are influencing this location.

SWP2b:



Yet another way to decearse dimensionality and make results more readable is Factor Analysis or Pricncipal component Analysis. For our survey results PCA was chosen over FA, due to several reasons. One of the main reason behind is that the survey was not constructed to test any theoretical model of latent variables. Despite it FA still was run over the data, to prove our point empirically. As the analysis of eigenvalues of the data showed 5 factors are used to replicate the original data. This is due to the fact that, these 5 factors are above 1 which means each of them can explain more variance than original variable. Although using so many factors, it only explains 46% of the cumulative variance which is not acceptable.



As this graph shows, there can be a significant logical link between the variables inside each of the factors. This proves there are no latent variables that should be taken into consideration.

Moving on to Principal Component Analysis