

Evaluation of the Social Climate by Poles - A Multidimensional Approach

Alicja Grześkowiak¹

¹Wrocław University of Economics

Abstract: Social climate is a complex concept which can be described by a range of factors. Its multidimensionality demands a specific approach in the analyses aimed at discovering patterns and relationships among its various aspects. The purpose of this paper is to analyze multifaceted opinions on various elements of the social climate in Poland in relation to socio-demographic characteristics as the age, gender, the labour market status, the level in the society. The principal components analysis combined with a clustering algorithm is applied to detect groups of persons evaluating the social climate similarly. The multifactor dimensionality reduction is used to search for interactions among elements described by categorical variables. All analyses are carried out on the base of survey data and are supported by appropriate visualization methods allowing the presentation of the results in a transparent manner. The calculations and the visualizations are carried out using suitable packages from the R program.

Keywords: social climate, multivariate analysis

1. Introduction

Measuring the social climate is a complex task as this phenomenon can be described by various factors. Several surveys focused exactly on this topic were conducted in the European Union countries under the *Eurobarometer* public opinion studies in 2009 [13], 2010 [12], 2011 [11], 2012 [10], 2013 [9] and 2014 [8]. These *Eurobarometer* Reports provide an overall social climate index calculated for the whole European Union and for the member countries separately. The comparison of this general index summarizing current situation in the European Union and in Poland within six consecutive years is presented in Fig.1.

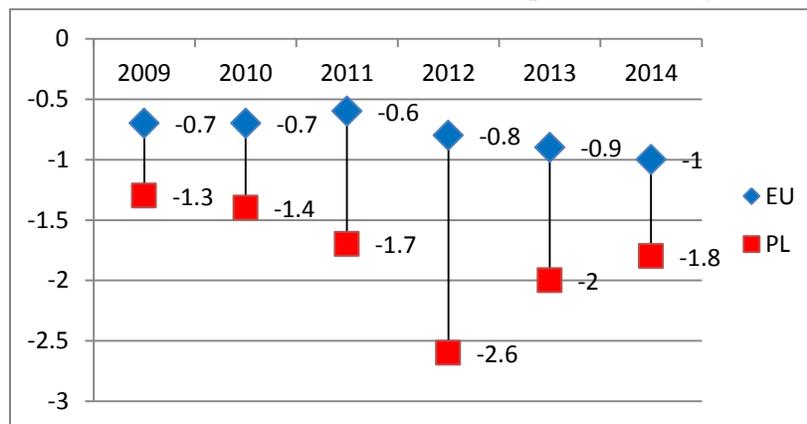


Fig. 1: Overall social climate index (current situation) in years 2009-2014 in the UE and in Poland (according to the *Special Eurobarometer 418 Report*, p.9)

In all the years under consideration, the values of the index were negative both for the European Union as a whole, and for Poland. A systematic gradual decline in the index value was recorded in the European Union from 2011, indicating a general deterioration of the situation in the last years. In Poland, by contrast, there has been a slow improvement after the minimum value in 2012. Nevertheless, the evaluation of the social climate in Poland was always much lower than the European average. This gives the rationale to a more thorough analysis of the phenomenon, taking into account various aspects of the assessment and their associations with socio-

economic variables. The main objective of this study is to analyze multifaceted opinions on various elements of the current social climate in Poland, by chosen multivariate techniques. Several specific objectives are also formulated. Firstly, the paper aims at the identification of interrelationships among social climate determinants. Secondly, the analysis is focused on finding clusters of individuals giving concordant opinions and to characterize them according to the socio-economic variables. Thirdly, the paper is focused on the recognition of the most influential social climate factors in respect to the satisfaction with life in general.

2. Data and Methods Description

This research is carried out on the base of the most recent individual data on the social climate taken from the *Special Eurobarometer 81.5* survey¹ which took place in 2014. The database for Poland contains 1082 records. The variables describing the social climate are divided into three areas: "personal circumstances, the national picture and social protection and inclusion" [8]. Each area is represented by a range of factors (Tab. 1).

TABLE I: Social Climate Components considered in the *Eurobarometer* survey

Social climate		
Personal circumstances	National Picture	Social protection and inclusion
Life in general	Cost of living	Healthcare system
Area where respondents live	Affordability of energy	Pension system
Personal job situation	Affordability of housing	Unemployment benefits
Financial situation of the household	Way public administration runs	Relations between people from different cultural or religious backgrounds or nationalities
	Economic situation	Way of addressing inequalities and poverty
	Employment situation	

Source: [8].

The answers given to various questions were used to construct composite indicators relating to these three areas. These indicators were constructed similarly to the social climate index, i.e. the answers very good/very satisfied were scored 10, rather good/fairly satisfied were scored 3,33, rather bad/not very satisfied were scored -3,33, very bad/not at all satisfied were scored -10 and "don't know" answers were not scored at all [8]. The final indicators for individuals were calculated as averages of the scores within a given area and denoted: personal, country and social, respectively.

The opinions of the social climate are considered in respect to some socio-economic characteristics represented by categorical variables available in the *Eurobarometer* survey, namely:

- age (intervals: 15-24; 25-39; 40-54; 55+),
- gender,
- labour market status (self-employed, employed, not working),
- level in the society - self assessment (low level, middle level, high level, refusal)
- age of leaving full-time education (15 or less, 16-19, 20+ and others including still studying, no full-time education, refusal, don't know).

According to the objectives of the study some multivariate techniques were applied to detect certain patterns in the data: the principal components analysis combined with the hierarchical clustering as well as the multifactor dimensionality reduction algorithm. The principal components analysis is used to replace original interrelated variables by a new set of uncorrelated variables ordered so that the first few represent the most of the total variation [4]. Various criteria are proposed to evaluate the importance of the components and to decide how many of them should be retained, i.e. the percentage of the explained variance, the eigenvalues greater than one, the scree plot examination and the interpretation usefulness [1].

The clustering methods allow searching for some regularities in the datasets. There are two main types of clustering algorithms: hierarchical and non-hierarchical. The hierarchical method used in this paper is the most

¹ European Commission, Brussels (2015): Eurobarometer 81.5 (2014). TNS Opinion [producer]. GESIS Data Archive, Cologne. ZA5929 Data file Version 2.0.0, doi:10.4232/1.12250

common approach giving a nested arrangement of objects in groups at each step of the procedure [3]. The hierarchical methods are based on the same general algorithm but there are various variants depending on the definition of the distance between two clusters [5]. In this paper the principal components analysis and the hierarchical agglomerative clustering procedure were combined. Firstly, the scores for the individuals were derived from the principal components and then the cluster analysis was performed on factor scores using the ade4TkGUI R package [7].

Another method used in this study is MDR (multifactor dimensionality reduction) proposed by Ritchie et al. [6] and described in detail in Hahn et al. [2] The method is used mainly in genetics, but its application may relate to any of the issues described by a set of non-metric variables among which there is a binary dependent variable describing the occurrence of the phenomenon. The procedure involves four steps [6]: the first is used to select K factors from all possible ones. In the second, all combinations of these factors' categories are expressed in the K-dimensional space. In the third, the ratio of the occurrences to the lack of occurrences of the phenomenon is calculated for each distinguished combination. The ratio greater than or equal to a defined threshold (e.g. 1) suggests that a set of factors' categories is favourable to the occurrence of the dependent variable. The fourth stage involves the estimation of the prediction/validation error, which enables the final selection of the model.

3. Social Climate Evaluation with Respect to Socio-Economic Characteristics

The analysis concerning the multidimensional social climate evaluation with respect to socio-economic characteristic was carried out in several stages. Firstly, the principal components analysis as an exploratory technique was performed for the social climate indicators. This allowed taking into account only two dimensions in the further analysis. Secondly, a hierarchical clustering Ward's method was applied to detect clusters of the respondents. Thirdly, the clusters were characterized by the socio-economic variables. The main results are visualized in Fig. 2-4 and are summarized in Tab.2.

The principal component analysis was carried out on the base of the correlation matrix. The scree plot illustrating the eigenvalues and the percentages corresponding to the proportion of the explained variance is presented on the left side in Fig 2. The first component reflects most of the variance i.e. 69,45%. The first and the second components together explain 90,75% of the total variance. Hence, two dimensions are sufficient for the further analysis. The variables factor map (Fig.2 on the right side) shows that the social and country indicators are highly positively correlated and almost uncorrelated with the personal one. The first dimension is related to all aspects while the second distinguishes between the personal indicator and the others.

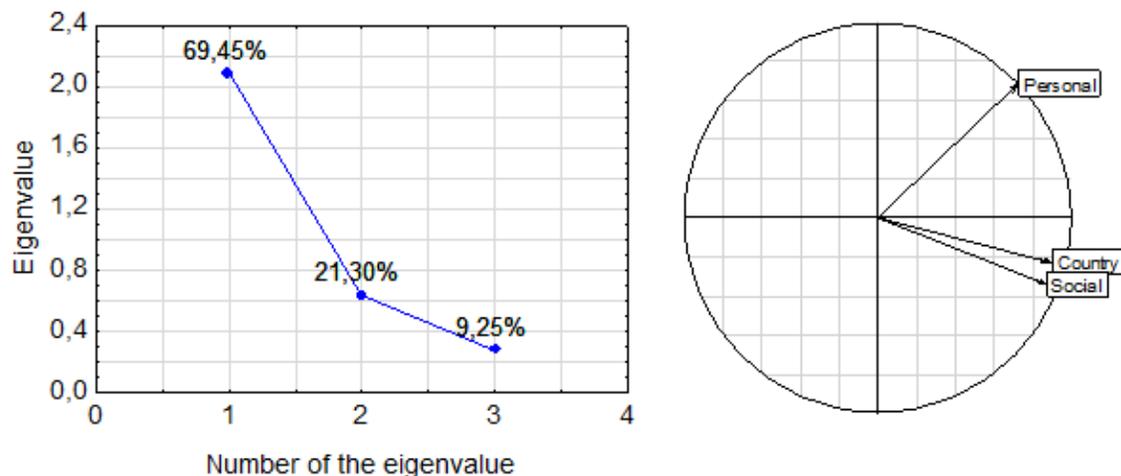


Fig. 2: PCA results: scree plot and variables factor map.

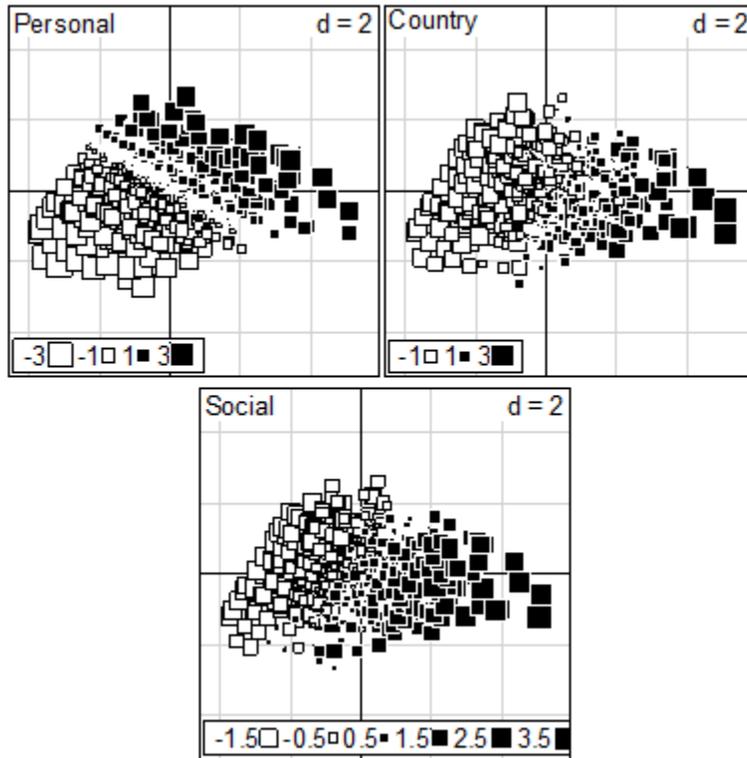


Fig. 3: Individuals represented on the factor map by squares reflecting the normalized values of particular variables.

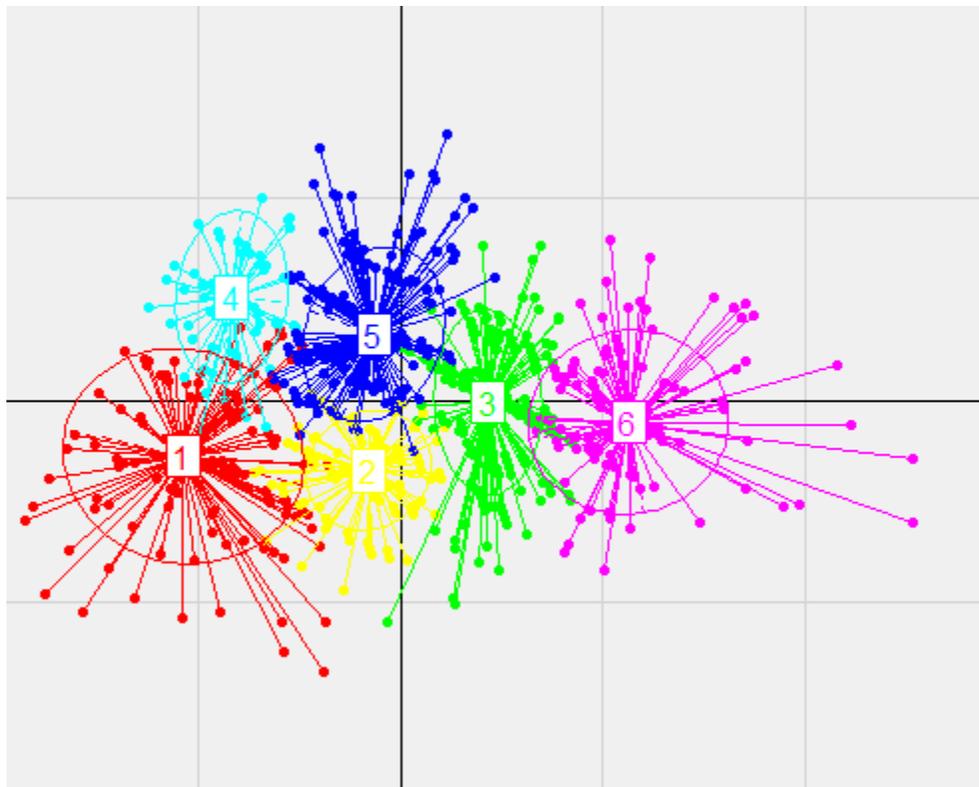


Fig. 4: Detected clusters marked on the factor map.

The individuals' scores projected on the plane spanned by the two first components are presented in Fig. 3. The size and the colour of the squares inform about the standardized values of the input indicators. As the social and country indicators are highly correlated the patterns for them are rather similar while the pattern for the personal indicator slightly correlated with the others is oriented differently.

The results of the clustering method plotted on the factor map are given in Fig. 4. The six identified clusters are marked by colours and numbers. The cluster characteristics are presented in Tab. 2.

TABLE II: Clusters characteristics (%)

Cluster	Gender	Age interval	Occupation	Level in society	Education
1	F: 64,6 M: 35,4	A: 5,0	A: 6,8	A: 45,3	A: 10,6
		B: 22,4	B: 35,4	B: 34,2	B: 53,4
		C: 22,4	C: 57,8	C: 11,2	C: 27,3
		D: 50,2		D: 9,3	D: 8,7
2	F: 55,1 M: 44,9	A: 11,4	A: 4,8	A: 38,9	A: 11,4
		B: 24,0	B: 30,5	B: 45,5	B: 49,1
		C: 19,8	C: 64,7	C: 12,0	C: 31,1
		D: 44,8		D: 3,6	D: 8,4
3	F: 56,3 M: 43,7	A: 16,5	A: 5,9	A: 18,4	A: 9,2
		B: 26,8	B: 45,6	B: 54,4	B: 40,8
		C: 19,5	C: 48,5	C: 24,6	C: 36,8
		D: 37,2		D: 2,6	D: 13,2
4	F: 55,9 M: 44,1	A: 11,8	A: 8,8	A: 20,6	A: 4,4
		B: 25,0	B: 42,6	B: 66,2	B: 44,1
		C: 26,5	C: 48,6	C: 11,8	C: 39,7
		D: 36,8		D: 1,4	D: 11,8
5	F: 63,0 M: 37,0	A: 13,2	A: 9,7	A: 22,2	A: 10,1
		B: 19,5	B: 41,6	B: 52,5	B: 39,3
		C: 24,5	C: 48,7	C: 21,4	C: 39,3
		D: 42,8		D: 3,9	D: 11,3
6	F: 54,6 M: 45,4	A: 19,7	A: 13,2	A: 10,5	A: 7,9
		B: 29,6	B: 41,4	B: 51,3	B: 24,3
		C: 23,7	C: 45,4	C: 32,9	C: 46,1
		D: 27,0		D: 5,3	D: 21,7

Note: Gender: F: female, M: male; Age interval: A: 15-24, B: 25-39; C: 40-54, D: 55+; Occupation: A: self-employed, B: employed, C: not working; Level in the society: A: low level, B: middle level, C: high level, D: refusal; Education (age of leaving): A: 15 or less, B: 16-19, C: 20+, D: others)

The cluster 1 consists of persons evaluating the social climate the worst. It has the highest percentage of women, the highest percentage of persons aged 55+ and the highest percentage of respondents assessing their level in the society as low. Its opposite is the cluster 6 consisting of persons evaluating the social climate the best in all dimensions. The specificity of this group is that the percentage of men and younger persons (up to 39) is the highest. The percentage of not working and poorly educated is the lowest. Individuals classified to this cluster evaluate their level in the society the best. The differences between other clusters are not so apparent but some regularities can be found. The cluster 3 can be called "moderate optimists" as the assessment of the social climate is quite high in all dimensions. In comparison to the cluster 6, its members are older, less educated, placed lower on the social ladder and less often undertaking work, especially in the form of self-employment. The individuals assigned to the clusters 4 and 5 evaluate the personal aspect relatively well and the other facets poorly (those belonging to the cluster 4 much worse). The members of the cluster 4 in comparison to those of the cluster 5 are younger, better educated and rarely describing their social level as high. The cluster 2 is characterized by moderate judgments on the social and country aspects and rather bad opinions on the personal

indicator. The percentage of not-working persons is the highest in this group. The individuals allocated to this cluster are relatively old and perceive their level in the society mostly as low or middle.

A general rule describing the results can be formulated. The high evaluation of the social climate, i.e. the position on the factor map more to the right, is associated with lower age, longer education, higher self-placement in the society and higher labour market activity. Nevertheless, there are some specific regularities resulting from the different perception of the social and country aspects as compared to the personal facet.

4. Overall Satisfaction with Life and Other Social Climate Determinants

This paragraph describes the identification of the most influential social climate factors in respect to the satisfaction with life in general. The data were recoded for the purpose of the multifactor dimensionality reduction. The overall satisfaction with life is treated as a dependent dichotomous variable with values 0 if a person is dissatisfied with life and 1 when satisfied. No-opinion cases were not taken into consideration. All other judgments on the current situation were defined as follows: 1 - bad, 2 - good, 3 - don't know. As the MDR technique demands equal frequencies of 0 and 1 cases of the dependent variable, all 251 "dissatisfied" records were used and 251 "satisfied" cases were randomly chosen from all "satisfied" ones.

Multifactor dimensionality reduction was performed using the package MDR from R program [14]. The procedure of the internal validation based on random division of the sample into three sets was applied: training, testing and validation ones [14]. Models of the number of factors not greater than three were considered (K=3). The results of the analysis are presented in Tab. 3 and Fig.5.

TABLE III: Multifactor dimensionality reduction results for K=3

Level	Best models	Accuracy		
		Training	Testing	Validation
1	FHH	77,90	71,35	76,62
2	PJS, FHH	77,90	73,10	77,76
3	ESC, PJS, FHH	82,06	75,04	80,19

Note: ESC - Economic situation in the country, FHH- financial situation of household, PJS - Personal job situation

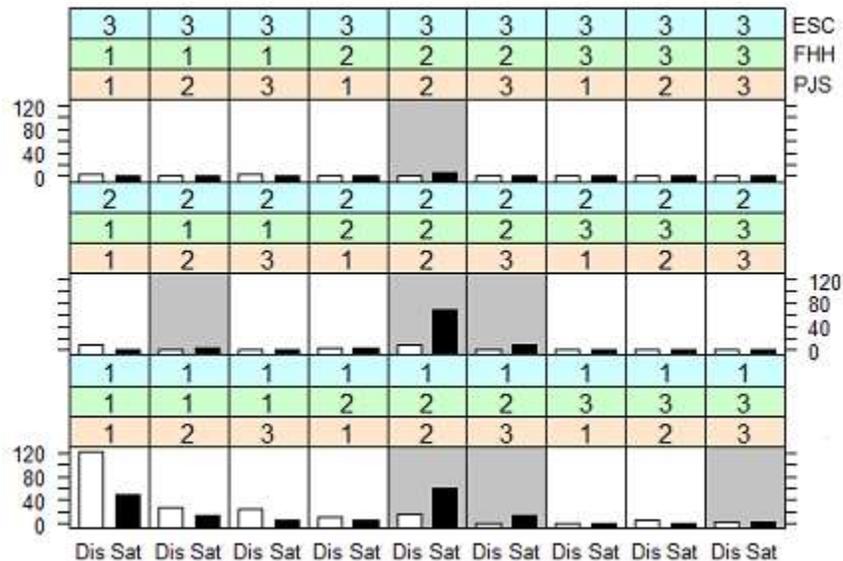


Fig. 5: Visualization of the multifactor dimensionality reduction results for the best model for K=3 (ESC, FHH, PJS - meaning as in Table 3, Sat-Satisfied, Dis-Dissatisfied).

The best model is a three-factor model consisting of the economic situation in the country, the financial situation of household and the personal job situation. The combination of these factors allows predicting the

satisfaction with life. The best model is illustrated in Fig.5. Numbers 1-3 represent the levels of factors taken into account, the black bars reflect the number of satisfied respondents and the white ones the number of dissatisfied respondents. The gray shading indicates the combinations of factors associated with the domination of the satisfied respondents.

The largest number of the satisfied Poles exceeding the dissatisfied ones occurs when all three recognized factors are rated "good". The second combination with an important excess of the satisfied over the dissatisfied persons consists of the responses "good" for the evaluation of the financial situation of household, "good" for the personal job situation and "bad" for the economic situation in the country. Other combinations characterized by the dominance of the satisfied respondents seem to have little significance. This distribution of answers suggests that the overall satisfaction with life is primarily evaluated through the prism of the personal factors related to the financial as well as professional situation and the perception of the economic situation in the country is of the secondary importance. Not surprisingly, the highest number of dissatisfied persons is associated with all three social climate categories rated as "bad".

5. Conclusions

The evaluation of the social climate in Poland, much lower than the EU average, indicates a negative perception of the situation in the country. Numerous measures are needed to improve various aspects of the economic and the social environment. The assessment of the social climate is not homogeneous. The differences among the groups evaluating the situation diversely are apparent. The socio-economic characteristics strongly differentiating opinions are: the age, the labour market status and the age of leaving full-time education. The evaluation of the social climate is also associated with the subjective perception of the level in the society. The assessment of the overall satisfaction with life is influenced mainly by factors related to the personal economic aspects: financial situation of household and personal job situation.

The multivariate data analysis techniques applied in this research allowed considering many dimensions of the social climate simultaneously. These methods are useful in analyzing multifaceted problems, including those described by categorical data, and also enable attractive visualizations of the results. Hence, they are very valuable tools in examining complex socio-economic phenomena.

6. Acknowledgements

This research was financed by Narodowe Centrum Nauki (National Science Centre) in Poland under the project entitled "*Non-metric multivariate data analysis as a tool for study of adults situation in the context of demographic changes*". Project number: 2012/05/B/HS4/02499.

This research uses Special Eurobarometer data: European Commission, Brussels (2015): Eurobarometer 81.5 (2014). TNS Opinion [producer]. GESIS Data Archive, Cologne. ZA5929 Data file Version 2.0.0, doi:10.4232/1.12250

7. References

- [1] D. J. Bartholomew, F. Steele, J. Galbraith, and I. Moustaki, *Analysis of Multivariate Social Science Data*, 2nd ed. Boca Raton: CRC Press, 2008, ch.5, pp. 117-144.
- [2] L. W. Hahn, M. D. Ritchie, and J. H. Moore, "Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions", *Bioinformatics*, vol. 19 no. 3, pp. 376-382, 2003.
<http://dx.doi.org/10.1093/bioinformatics/btf869>
- [3] J. Jobson, *Applied Multivariate Data Analysis: Volume II: Categorical and Multivariate Methods*. New York: Springer Science & Business Media, 1992, ch.10.2.1, pp. 519-534.
- [4] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York: John Wiley & Sons, Inc, 2002, ch.1, pp.1-9.
- [5] A. C. Rencher, *Methods of Multivariate Analysis*, 2nd ed. USA: John Wiley & Sons, Inc, 2002, ch.14, pp.451-503.
<http://dx.doi.org/10.1002/0471271357.ch14>

- [6] M. D. Ritchie, L. W. Hahn, N. Roodi, L. R. Bailey, W. D. Dupont, F. F. Parl, and J. H. Moore, "Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer", *The American Journal of Human Genetics*, 69, pp. 138-147, 2001.
<http://dx.doi.org/10.1086/321276>
- [7] J. Thioulouse and S. Dray, "Interactive multivariate data analysis in R with the ade4 and ade4TkGUI packages", *Journal of Statistical Software*, vol. 22, issue 5, pp. 1-14, September 2007.
- [8] TNS Opinion & Social. (November 2014). Special Eurobarometer 418: Social Climate. Report. Available: http://ec.europa.eu/public_opinion/archives/ebs/ebs_418_en.pdf
- [9] TNS Opinion & Social. (October 2013). Special Eurobarometer 408: Social Climate. Report. Available: http://ec.europa.eu/public_opinion/archives/ebs/ebs_408_en.pdf
- [10] TNS Opinion & Social. (September 2012). Special Eurobarometer 391: Social Climate. Report. Available: http://ec.europa.eu/public_opinion/archives/ebs/ebs_391_en.pdf
- [11] TNS Opinion & Social. (October 2011). Special Eurobarometer 370: Social Climate. Report. Available: http://ec.europa.eu/public_opinion/archives/ebs/ebs_370_en.pdf
- [12] TNS Opinion & Social. (October 2011). Special Eurobarometer 349: Social Climate. Report. Available: http://ec.europa.eu/public_opinion/archives/ebs/ebs_349_en.pdf
- [13] TNS Opinion & Social. (January 2010). Special Eurobarometer 315: Social Climate. Full Report. Available: http://ec.europa.eu/public_opinion/archives/ebs/ebs_408_en.pdf
- [14] S. J. Winham and A. A. Motsinger-Reif "An R package implementation of multifactor dimensionality reduction", *BioData mining*, 4: 24, 2011.
<http://dx.doi.org/10.1186/1756-0381-4-24>