

JOURNAL OF SCIENCE



SAKARYA UNIVERSITY

Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University
<http://www.saujs.sakarya.edu.tr/en/>

Title: The Classification of OECD Countries in Terms of Life Satisfaction Using Partial Least Squares Discriminant Analysis

Authors: Esra Polat

Received: 2019-10-14 14:36:39

Accepted: 2020-01-28 13:47:04

Article Type: Research Article

Volume: 24

Issue: 2

Month: April

Year: 2020

Pages: 365-376

How to cite

Esra Polat; (2020), The Classification of OECD Countries in Terms of Life Satisfaction Using Partial Least Squares Discriminant Analysis. Sakarya University Journal of Science, 24(2), 365-376, DOI:

<https://doi.org/10.16984/saufenbilder.632820>

Access link

<http://www.saujs.sakarya.edu.tr/tr/issue/52471/632820>

New submission to SAUJS

<http://dergipark.org.tr/en/journal/1115/submission/step/manuscript/new>



The Classification of OECD Countries in Terms of Life Satisfaction Using Partial Least Squares Discriminant Analysis

Esra POLAT¹

Abstract

Life satisfaction (LS) measures how people assess their lives as a whole, not their present emotions. Measuring emotions can be very subjective, but it is still a useful completion to more objective data when comparing quality of life across countries. Many questionnaires are used to measure especially LS and happiness. The Partial Least Squares Discriminant Analysis (PLSDA) is a statistical method for classification and includes an ordinary Partial Least Squares Regression, where the dependent variable is categorical that represents each observation's class membership. In this study, the purpose is to classify 35 OECD countries correctly to their predefined classes (above or below the average LS level of OECD) by using year 2017 Better Life Index data. In the analyses PLSDA, a flexible supervised classification method, is used. PLSDA is a preferable alternative method in case of some assumptions not satisfied for classical discriminant analysis. The results showed that PLSDA has a satisfying classification performance and self-reported health (SH) is only effective variable in determining the LS levels of countries.

Keywords: Better Life Index, classification, life satisfaction, OECD countries, Partial Least Squares Discriminant Analysis

1. INTRODUCTION

There is much more than Gross Domestic Product (GDP) numbers and economic statistics. Therefore, the current economic and financial crisis has refocused interest in other factors. The trick is to decide what works for a better life and the way of measuring progress. In most OECD countries, inequality is broadening and more

money does not make people feeling better. So that what else should be measured that thought to affect the life happiness? The OECD interested on this question over ten years ago; and work such as the Stiglitz-Sen-Fitoussi Commission in France, and in recent times national attempts such as the UK's programme Measuring National Well-Being [1].

¹ Corresponding Author: espolat@hacettepe.edu.tr

Hacettepe University, Faculty of Science, Department of Statistics, ORCID ID: 0000-0001-9271-485X

In 2011, the OECD Better Life Initiative took an interactive step with the Better Life Index, an international comparable well-being indicator (How's Life? Report) and its interactive tool, the Better Life Index (BLI) inviting users to look at how their country measures up on the subjects that significant for them, a wide range of topics from education to air pollution, from health to income etc. [1].

A key reason for measuring well-being is to understand whether, where and how life is getting better for people. In "How's Life? 2017: Measuring Well-being" report provides an overview of OECD countries' achievements across 11 dimensions of current well-being and four different "capital stocks" that help to sustain well-being overtime. It features a various set of statistics, ranging from household wealth to times spent on leisure, and from air pollution to how safe people feel walking alone at night. Since the last 10 years have been a turbulent time in most OECD economies, the particular focus on changes in people's well-being. It seeks to address the simple question: Is life today better or worse than it was in 2005, before the financial crisis took hold? [2].

Life satisfaction (LS) is highest in Denmark, Norway and other Scandinavian countries, also in Switzerland, New Zealand, Canada and Australia - countries with high levels of employment, quality of jobs and population health [3]. In countries with the lowest LS, employment levels and, usually, life expectancy are below the OECD average. The BLI indicates that having strong relationships with friends and deriving pleasure from a good work-life balance and personal safety is associated with high LS. Scandinavian countries score high in these areas, also in Spain where data that underpin BLI indicate that 96% of people know someone they can trust in the moment of need, one of the OECD's highest rates [3].

The BLI shows that there is little difference in LS between men and women in 35 countries. Nevertheless, in OECD countries individuals with bachelor's degree have tendency to have higher LS than those who are merely primary school

graduated. BLI includes 35 countries and measures well-being over 11 dimensions (civic engagement, community, education, environment, health, housing, income, jobs, life satisfaction, safety, work-life balance). This index clears that if a country has a good performance in economic, this means not directly it will indicate same good performance in terms of well-being. For example, Mexico and Turkey indicate a good performance in some form of civic engagement. South Africa scores inadequately compared to rich countries in terms of many indicators, however, it has a relatively strong public conscience and work-life balance [3].

Personal security is also a problem in some rich countries. BLI shows that in Australia, New Zealand and US people feel unsafety in high level. The countries who perform well above the average in work-life balance have various economic levels such as Hungary, Ireland, Italy and Russia. Estonia, Germany, Japan, Korea and Poland are among the countries with the best general education and skill levels. Decent housing is a significant component of well-being. BLI reveals that good housing conditions are often connected with good economic outcomes. Canadians and Americans demonstrate a tendency to benefit from best housing circumstances [3].

BLI's online interactive tool also permits users to directly tell what is substantial to their own well-being. Up to now, in 180 countries this tool has been used by more than 110000 people. In general, online users rank education, health and LS as the most essential elements for their well-being. The education is commonly considered to be the most important of the 11 well-being dimensions in Latin America. However, regional differences appear. LS and work-life balance are among the top precedencies in North America even though community environment and health are the base interests in Europe [3].

LS measures how people evaluate their lives as a whole, not their present emotions. Considering the all OECD countries using a scale between 0 and 10, people gave an average score of 6.5 for scoring their overall LS. However, LS is not

shared equally in the OECD. In some countries such as Portugal, Greece, Turkey - with an average of 5.5 scores or less, seen relatively low level of overall LS. At the other end of the scale, scores reach 7.5 in Denmark, Iceland, Finland, Switzerland and Norway [4].

The term of a better life first appeared in the 2000s. In 2001, the OECD published a report on the better lives of countries [5]. Osberg and Sharpe [6] developed an economic well-being index by merely in view of economic variables for chosen OECD countries: Australia, Canada Sweden, Norway, U.K, U.S. Kerényi [7] also studied on and introduced a BLI for countries. Kasparian and Rolland [8] developed a better quality of life index based on diverse data from OECD countries. Stevenson and Wolfers [9] interested on the relationship between income and well-being so that evaluations made on LS. Mizobuchi [10] suggested a combined indicator of overall well-being, to measure the performance of each country in supplying well-being to its citizens. He applied Data Envelopment Analysis to form a group of 11 separate well-being indicators into a combined indicator, using the World Bank's production base estimates for each country. Akar [11] assessed BLI as an alternative tool for measuring well-being for Turkey. As a result of this it is found that the lowest BLI value belongs to Turkey among OECD countries. Durand [12] discussed the advantages and disadvantages of several approaches for introducing and spreading information on multidimensional well-being to different people, containing the OECD BLI. The progress made in developing measures of well-being is exemplified and the statistical agenda for improving present indicators and building-up new ones is outlined [5]. Gundogan Aşık and Altın Yavuz [5] compared six different methods for modeling the LS using the OECD BLI Data. They have found that if solely classification is interested, the robust discriminant analysis can be used for modeling of LS. But robust logistic ridge regression could be used in case of determining the effective levels for LS. Başol [13] aimed to discover the dynamics affecting the LS in OECD countries by using the year 2016 BLI data. The results of this research using structural equation modeling technique and

model development strategy showed; health and positive work quality positively affect LS; income and negative job quality negatively affect LS.

PLSDA is a good alternative to classical discriminant analysis, since in some circumstances classical one could not be used while PLSDA could be performed; for instance, the situations such as the number of explanatory variables exceeding the number of observations ($p \gg n$). Moreover, it can be performed on data even if in case of missing values and multicollinearity problem and non-normality [14]. The purpose of this study is classify 35 OECD countries by using PLSDA according to their LS level (above or below the average score of 6.5 across the OECD) by using the potential effective 23 variables on LS that constituting quality of life.

2. PARTIAL LEAST SQUARE DISCRIMINANT ANALYSIS

PLSDA is a supervised classification method, since it must have primary information about the class memberships of the samples. Barker and Rayens [15] compared PLSDA with Linear Discriminant Analysis (LDA) and mentioned that PLSDA has advantages over classical Discriminant Analysis (DA) such as choosing of variables and reduction of noise [16].

The PLSDA method use the same algorithm for Partial Least Squares Regression (PLSR), the only difference is Y has discrete values used for showing class memberships of each observation. PLSR searches for a direct relationship between dependent variable and the explanatory variables. X matrix with $n \times p$ dimension shows the independent variables; n is the number of observations and p is the number of explanatory variables. Y matrix with $n \times q$ dimension corresponding to the dependent variable. Here, q represents the number of dependent variables. X and Y are decomposed by scores (or components, or latent variables) and the size of the data is reduced as shown in Eq. (1) and Eq. (2) [16].

$$X = TP' + E \quad (1)$$

$$Y = UQ' + F \tag{2}$$

T and U are the score matrices for X and Y, respectively; P shows the loading matrix of X; E is the error term for X; Q is the loading matrix for Y and F is the residual matrix for Y. In PLSR while choosing components, their relationship with Y is also considered different from Principal Component Regression (PCR). The optimal number of components are generally lower than PCR. T components are orthogonal and they are estimated as in Eq. (3) by using W^* , the weight matrix [16].

$$T = XW^* \tag{3}$$

The T components are good predictors of Y and the PLSDA model is written as in Eq. (4). F shows the deviations between the real and predicted response.

$$Y = TQ' + F \tag{4}$$

Inserting Eq. (3) in Eq. (4) the model can be updated lastly as in Eq. (5) and turns to a regression model as in Eq. (6).

$$Y = XW^*Q' + F \tag{5}$$

$$Y = X\beta + F \tag{6}$$

The regression coefficients are obtained as $\beta = W^*Q'$, where W^* can be obtained as in Eq. (7). W is defined using a set of weighting loadings, which maximizes the covariance between X and Y [17]. The detailed information about PLSR model and its classical algorithm's steps could be found in Wold et al. [18].

$$W^* = W(PW)^{-1} \tag{7}$$

In case of two classes in the data set, in PLSDA the matrix Y is coded to 0 or 1 (G=2). In case of multiple classes (G>2), several models could be constructed with 0 and 1 encoding, or the PLS2 algorithm is used by constructing a matrix (nxG), in which each column shows a class [16, 19]. An important stage of constructing a PLSDA model is the determination of the ideal number of LVs.

For this purpose, usually cross-validation (CV) is used. In this method the data set is divided by training samples and validation samples and the models are built with the separated observations for validation sample. The prediction errors are computed for separated samples using various numbers of LVs. The process is repeated until all samples are predicted. The PLSDA model gives a number by using Eq. (5), not reading completely 0 or 1. Hence, constructing threshold values is necessary for defining the class limits. The threshold is estimated by using Bayesian theorem in many approaches [19] or by constructing confidence limits for each classified object. Usually resampling techniques such as bootstrap could be used for the calculation of these confidence intervals [16].

3. CLASSIFICATION PERFORMANCE MEASURES

The calculated classification measures are described in Ballabio et al. [20]. These measures are used to assess the performance of classification methods such as classical DA or PLSDA. The classification results can be showed in confusion matrix (or contingency table). Since G represents the number of classes, the **confusion matrix** dimension is G x G. It could be showed as in Table 1. Each element c_{gk} shows the number of samples belonging to class g and assigned to class k. Hence, the diagonal elements c_{gg} show the number of correctly assigned observations, while off-diagonal elements show the numbers of unclassified observations [20, 21].

Table 1. Representation of confusion matrix (in case of G classes)

		Assigned Class				
		1	2	...	G	
True Class	1	c_{11}	c_{12}	...	c_{1G}	n_1
	2	c_{21}	c_{22}	...	c_{2G}	n_2

	G	c_{G1}	c_{G2}	...	c_{GG}	n_G
		n'_1	n'_2		n'_G	n

Three popular class-based measures (sensitivity, precision and specificity) are used for estimating the classification performance obtained on each class. They are computed on each class individually and show different sides of the classification [20].

Sensitivity defining the model ability of correctly recognizing samples of the g -th class and is given as in Eq. (8) [20, 21]:

$$Sn_g = \frac{c_{gg}}{n_g} \quad (8)$$

Here, c_{gg} (the diagonal elements of confusion matrix) showing the correctly classified samples, n_g is the total number of objects that member of the g -th class. In case of all the samples that member of the g -th class are correctly assigned ($c_{gg} = n_g$), Sn_g equals to 1. The unassigned objects of the g -th class are not taken under consideration for the sensitivity computation.

Precision represents the capability of a classification model not containing objects of other classes in the examined class. It shows the ability of a classifier avoiding wrong predictions in that class and given by Eq. (9) [20, 21]:

$$P_{rg} = \frac{c_{gg}}{n'_g} \quad (9)$$

Here, the total number of objects assigned to the g -th class showed by n'_g . If all the objects assigned to class g correspond to the samples member of class g , P_{rg} equals to 1.

Specificity characterizes the capability of a classifier to reject the samples of all the other classes and is given as in Eq. (10) [20, 21]:

$$Sp_g = \frac{\sum_{k=1}^G (n'_k - c_{gk})}{n - n_g} \quad \text{for } k \neq g \quad (10)$$

Each element of confusion matrix c_{gk} represents the number of objects belonging to class g and

assigned to class k . Hence, n'_k shows the total number of objects classified to the k -th class:

$$n'_k = \sum_{g=1}^G c_{gk}$$

This measure computed as the ratio of “samples not member of the g -th class also not assigned to the g -th class“ over “the total number of samples not member of the g -th class ($n-n_g$)”. Sp_g equals to 1, in case of the objects not member of class g are never classified to g . Not classified objects are not taken under consideration for the specificity computation.

Until now the three measures that we examined give the classifier performances on each specific class, however, they do not yield total assessment of the classification quality. Hence, by clustering class measures in different ways, global measures of classification performances are computed.

Accuracy (AC) is another index helps for evaluating the classification quality. It is also named as overall agreement/predictive ability/classification rate/success rate, total accuracy. It is given as in Eq. (11) and shows the ratio of correctly classified objects. It takes the values between 0 (no correctly classified objects) to 1 (perfect classification) [20, 21]:

$$AC = \frac{\sum_{g=1}^G c_{gg}}{n} \quad (11)$$

n is the total number of samples and not classified objects are not used for the accuracy computation. “Misclassification error” is the complementary of it and defined as the ratio of objects classified to a wrong class.

Non error rate (NER) is the mean of the class sensitivities [20]:

$$NER = \frac{\sum_{g=1}^G Sn_g}{G} \quad (12)$$

Error Rate (ER) is given as: $ER = 1 - NER$, using the non-error rate.

Ratio of not assigned samples is the fraction of the objects that could not have assigned in the modelled classes. Not assigned samples are not used in the specificity, sensitivity, non error rate and error rate computations [20].

Classification results could be presented by using graphs such as **ROC (Receiver Operating Characteristics) curves**. In Classification toolbox for MATLAB [22], as a result of PLSDA these curves are also given.

A ROC curve is a plot of sensitivity and specificity, used for classification studies of two class data sets. Its discrimination threshold is not fixed. By using contingency table, a single value of sensitivity and specificity can be computed. So that each contingency table shows one point in the ROC space. For each threshold value, a classification rule is computed and the related contingency table is obtained. The best possible classification method would produce a point in the upper left corner of the ROC space, standing for maximum sensitivity and specificity, however, a random classification yields points along the diagonal line from the left bottom to the top right corners. ROC curves are computed for each class, separately, by changing the threshold of assignments. The area under the ROC curve (AUC) can be used as estimator of the class discrimination; it is shown in the plot title for each class [20]. ROC is a probability curve and AUC shows degree or measure of separability. It shows model ability of distinguishing between classes.

Formula and extra details on these classification measures are given in html help files created for the Classification Toolbox for MATLAB [22, 23].

4. APPLICATION AND RESULTS

The Better Life Index is applied to 35 OECD members. These OECD countries are; Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia,

Spain, Sweden, Switzerland, Turkey, United Kingdom, United States [2, 5].

The data set is obtained from OECD 2017 BLI data. LS used as the dependent variable and coded as two classes for each of 35 OECD countries. 6.5 is the average LS score across the OECD.

$$y_i = \begin{cases} 0, & \text{Countries with a LS score less than 6.5 / low LS level} \\ 1, & \text{Countries with a LS score greater than 6.5 / high LS level} \end{cases}$$

The independent variables are 23 sub-dimensions given in OECD BLI data set. 10 dimensions and their sub-dimensions can be examined in Table 2. The analysis made by using Classification Toolbox (MATLAB) and XLSTAT (Excel) programs. XLSTAT running on Excel and lets users for analyzing, customizing and sharing their results within Microsoft Excel [22, 23, 24].

Table 2. OECD Better Life Index 2017 dimensions and sub-dimensions

Dimensions	Sub-Dimensions
Housing	<ul style="list-style-type: none"> Dwellings without basic facilities (DW) Housing expenditure (HE) Rooms per person (RP)
Income	<ul style="list-style-type: none"> Household net adjusted disposable income (HI) Household net financial wealth (HFW)
Jobs	<ul style="list-style-type: none"> Labour market insecurity (LM) Employment rate (ER) Long-term unemployment rate (LUR) Personal earnings (PE)
Community	<ul style="list-style-type: none"> Quality of support network (QN)
Education	<ul style="list-style-type: none"> Educational attainment (EA) Student skills (SS) Years in education (YE)
Environment	<ul style="list-style-type: none"> Air pollution (AP) Water quality (WQ)
Civic engagement	<ul style="list-style-type: none"> Stakeholder engagement for developing regulations (SE) Voter turnout (VT)
Health	<ul style="list-style-type: none"> Life expectancy (LE) Self-reported health (SH)
Safety	<ul style="list-style-type: none"> Feeling safe walking alone at night (SWA) Homicide rate (HR)
Work-Life Balance	<ul style="list-style-type: none"> Employees working very long hours (EH) Time devoted to leisure and personal care (TLP)

Resource: <https://stats.oecd.org/Index.aspx?DataSetCode=BLI>

The data set could be analyzed by using classical DA. However, there are several assumptions for DA that must be checked. Firstly, an assumption that the discriminating variables follow the multivariate normality must be checked.

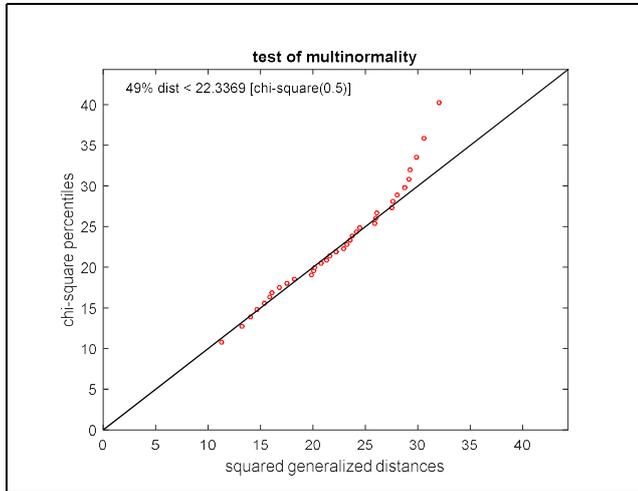


Figure 1. The chi-square Q-Q plot

Multivariate normality test has been implemented by using Mahalanobis Distances. Figure 1 shows that there are some deviations from the straight line, therefore, possible deviations from a multivariate normal distribution. It can be concluded that, this data set not meets multivariate normality assumption since the plot indicates departures from multivariate normal distribution explicitly.

Multicollinearity exists in any model when two or more independent variables in the model are related to each other. There are several different numerical methods for exploring multicollinearity connections. VIF and Tolerance values are statistics which the researches usually prefer. The multicollinearity detection is done by using Microsoft Excel XLSTAT program. In practice, if any of the VIF values is equal or larger than 10, there is a near collinearity. From Table 3 it is clear that the VIF values for RP, HI, LM, ER, LUR, PE, SS, WQ are 13.398, 17.475, 11.142, 13.995, 13.552, 22.475, 11.112, 18.865, respectively. Hence, there is a multicollinearity problem for this dataset.

Table 3. Multicollinearity statistics result

Statistic	DW	HE	RP	HI
Tolerance	0.136	0.340	0.075	0.057
VIF	7.379	2.937	13.398	17.475
Statistic	HFW	LM	ER	LUR
Tolerance	0.189	0.090	0.071	0.074
VIF	5.295	11.142	13.995	13.552
Statistic	PE	QN	EA	SS
Tolerance	0.044	0.105	0.142	0.090
VIF	22.475	9.556	7.054	11.112
Statistic	YE	AP	WQ	SE
Tolerance	0.350	0.102	0.053	0.294
VIF	2.856	9.769	18.865	3.400
Statistic	VT	LE	SH	SWA
Tolerance	0.295	0.163	0.116	0.175
VIF	3.385	6.123	8.600	5.711
Statistic	HR	EH	TLP	
Tolerance	0.166	0.173	0.180	
VIF	6.015	5.774	5.549	

Since multivariate normality assumption and independence among predictors are not satisfied, instead of classical DA, PLSDA could be implemented on the data set. Before it is mentioned that PLSDA is not affected by these assumptions. Firstly, for PLSDA the ideal number of LVs must be determined. For this purpose, error rate CV against number of LVs graph could be used [23].

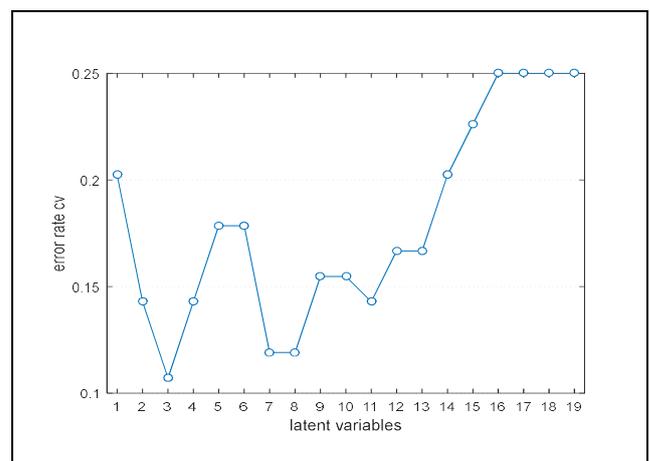


Figure 2. Error rate in CV versus number of components in PLSDA. CV was implemented 5 groups, obtained by venetian blinds method

From Figure 2, it is clear that by choosing 3 LVs much simple classification model is computed and CV error is very low for 3 LVs. After the ideal number of LVs is determined, PLSDA model can be obtained on the training samples. The final PLSDA model is obtained by choosing 3 LVs and 5 CV groups for validation. The confusion matrices constructed for fitting and CV as given in Table 4. The outputs of the classification model are collected in confusion matrix and it is the preliminary stage of evaluating the classification performance. The last column of the Table shows the number of unclassified objects for each class. From Table 4, it can be observed that, in fitting, 1 out of 21 high samples are classified into low class (3 in CV), 1 out of 14 low samples are classified into high class (1 in CV), and finally there is not any samples that are not assigned. It is clear from Table 4 that %94.29 of the countries are correctly classified for the training sample. For prediction sample it is seen from Table 4 that %88.57 of the countries are correctly classified.

Table 4. Confusion matrices for fitting and CV with 5 venetian blinds groups

Real Class	Predicted Class		
	High	Low	Not assigned
Fitting			
High	20	1	0
Low	1	13	0
Cross validation			
High	18	3	0
Low	1	13	0

Table 4 reveals that in case of two classes the confusion matrix could be shown as in below (in which the high class is defined as positive, P and low class as negative, N):

		Predicted Class	
		High (P)	Low (N)
Real Class	High (P)	TP	FN
	Low (N)	FP	TN

Here, TP (True Positive) represents the number of high objects correctly classified as high, TN (True Negative) shows the number of low objects correctly classified as low, FN (False Negative) represents the number of high objects wrongly classified as low and FP (False Positive) shows

the number of low objects wrongly classified as high.

TP/(TP+FN) is used for computing the high class sensitivity. Class sensitivity values range between 0 and 1 and defining the model capability to correctly distinguish objects that are member of that class. For instance, if not any of the high samples are assigned to low class (FN equals to 0), the sensitivity for high class can be equal to 1.

On the contrary, TN/(FP+TN) is used for high class specificity. The class specificity values range between 0 and 1 and defining the model capability of rejecting objects of all other classes. For instance, if not any of the low samples are assigned to high class (FP equals to 0), the specificity for high class can be equal to 1 [23].

In Table 5, the classification performance measures of PLSDA model are presented. It is known that when there are only two classes, sensitivity and specificity of two classes are symmetrical, as a result, all the time sensitivity of the high class equals to specificity of the low class and vice versa. The results of fitting (in case of all training samples used for modelling) shows that the low class's specificity and sensitivity values are 0.95 and 0.93, respectively. Taking under consideration only the classified samples, this means that 93 % of the low training samples (13 out of 14) are correctly classified as low and 95 % of the high training samples (20 out of 21) are correctly classified as high. Because of sensitivity and specificity values show similarity, it can be concluded that the type of error is balanced, hence, there is not special trend in the model for recognizing high samples as low, or vice versa. The model NER and ER in fitting are equal to 0.94 and 0.06, respectively. Finally, the classification performance of CV can be compared with model fitting. Cross-validated and fitting results are more similar for high class and less similar for low class. Although small difference can be seen, still it could be concluded that the PLSDA classification model can be supposed to be reliable and stable, since the classification performance is not badly affected by samples taken out from the training set during the CV procedure.

Table 5. Classification performance measures for both fitting and CV (for 5 venetian blinds groups)

	High		
	sensitivity	specificity	precision
Fitting	0.95	0.93	0.95
Cross validation	0.86	0.93	0.95
	Low		
	sensitivity	specificity	precision
Fitting	0.93	0.95	0.93
Cross validation	0.93	0.86	0.81
	NER	ER	AC
Fitting	0.94	0.06	0.94
Cross validation	0.89	0.11	0.89

The evaluation of the classification performance of a model could also be made by using ROC curves. A perfect model's AUC will be close to 1 means that having a good measure of separability. A poor model's AUC will be close to 0 meaning that having the worst measure of separability. This kind of poor model predicts 0s as 1s and 1s as 0s (means lows as highs, highs as low). Moreover, in case of AUC is 0.5, it refers to model's incapability of separation. In Figure 3, ROC curves for high (upper) and low (lower) classes are shown. Table 5 also reveals that the ROC curves of both classes are nearly perfect. The plots on the right of Figure 3 are ROC curves, as showing the sensitivity and specificity values as the class threshold for assigning samples to the class is changed. The class threshold is chosen at the point where the number of FPs and number of FNs is minimized and hence, its value corresponds to the point where the specificity line crosses the sensitivity line.

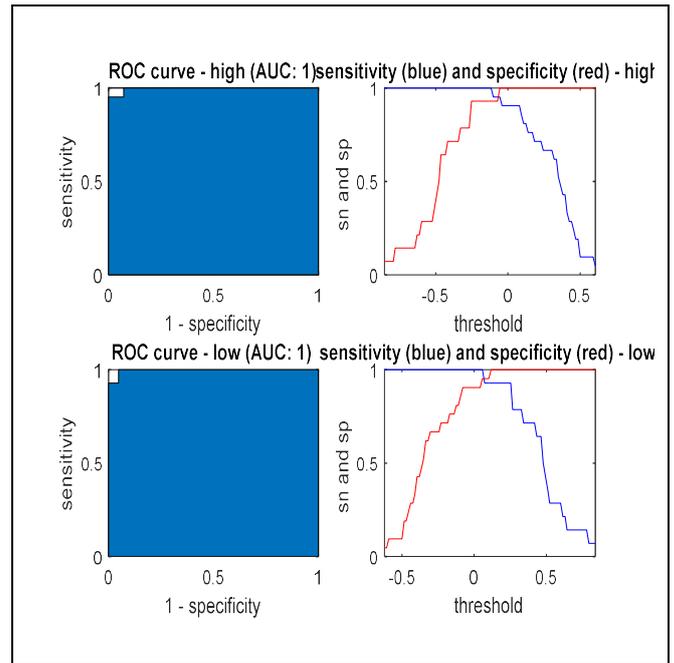


Figure 3. ROC curves (LEFT SIDE) and plots of specificity (red) sensitivity (blue) values (RIGHT SIDE)

Since PLSDA's origin comes from PLSR algorithm, the regression coefficients of the variables are obtained using this algorithm. The significant variables, for classifying objects to their correct classes, will have positive coefficients that contributing in increment of the class calculated response. Until now all results are obtained in Classification Toolbox, but in order to see which variable/variables are important in discriminating the classes, the results are obtained in XLSTAT.

The standardized coefficients are given in Table 6. It can be used to compare the relative weight of the variables in the model. For the computation of confidence intervals of coefficients, PLSR do not use the classical formulae based on the normality hypotheses used in Ordinary Least Squares regression. A bootstrap method gives confidence intervals estimations. If the absolute value of a coefficient is higher, weight of the variable in the model is also higher. In case of interval estimation of standardized coefficients contains 0, the weight of the variable in the model is unimportant [14].

Table 6. Standardized coefficients of the model

Variable	Coeff.	Std. Dev.	LB (95%)	UB(95%)
DW	-0.078	0.202	-0.489	0.332
HE	0.092	0.184	-0.281	0.465
RP	0.173	0.157	-0.147	0.492
HI	0.036	0.105	-0.178	0.250
HFV	0.003	0.143	-0.289	0.294
LM	-0.239	0.208	-0.661	0.184
ER	0.119	0.141	-0.168	0.406
LUR	-0.184	0.156	-0.502	0.133
PE	0.131	0.079	-0.029	0.292
QN	-0.007	0.121	-0.253	0.239
EA	-0.025	0.188	-0.406	0.356
SS	-0.290	0.162	-0.619	0.039
YE	0.084	0.159	-0.240	0.408
AP	0.008	0.172	-0.340	0.357
WQ	0.021	0.114	-0.211	0.253
SE	-0.158	0.251	-0.668	0.352
VT	0.051	0.202	-0.358	0.461
LE	-0.011	0.157	-0.329	0.308
SH	0.394	0.168	0.052	0.736
SWA	-0.030	0.125	-0.283	0.224
HR	0.236	0.173	-0.116	0.589
EH	-0.104	0.206	-0.523	0.316
TLP	0.020	0.178	-0.342	0.382

The results in Table 6 indicates that the only important variable that determine the statuses of high and low LS levels of OECD countries is "self-reported health (SH)" that means percentage of people whose feeling healthy is the most important determinant of LS.

5. CONCLUSION

Quality of life of countries mainly can be understood from LS variable. The countries policies for making progress about their economic prosperity will be inevitably affected by researches on influences of other variables of quality of life on LS. Here, a significant variable LS that effects the quality of life is taken under consideration. A comparison between the welfare levels of countries in terms of many different areas can be derived by LS. Particularly in these days, the welfare of the countries identified

merely by the income does not show that the welfare of the country is well. There are many different determinants of LS, that OECD surveys sum up them under titles like income, housing, jobs, education, community, environment, health, civic engagement, work-life balance and safety. Different from previous studies on this field, in this study, the sub-dimensions under these variables are used to find the most effective variable in determining the countries' LS levels. The analyzes showed that data set is non-normal and also there is a multicollinearity problem. Therefore, classical DA couldn't be used and PLSDA is preferred. As a result of PLSDA %94.29 of the countries are correctly classified for the training sample and %88.57 of the countries are correctly classified for validation sample. PLSDA has a good classification performance. Moreover, it is found that self-reported health (SH) is the only important variable in determining life satisfaction levels of 35 OECD countries.

Systematic health surveys are done by most of OECD countries for enabling participants to report on various statuses of their health. "How is your health?" is a frequently asked question for collecting data about self-perception health status. In spite of this question is non-objective, the responses are used to be a well estimator of people's future healthcare. In the OECD, around 69% of the adult population tell their health is "good" or "very good". Although 88% of adults say their health is "good" in New Zealand and Canada, less than 40% of people express their health as "good" or "very good" in Korea and Japan. Cultural, regional or other elements could affect the answers of this popular health question.

Men most likely to report better health compared to women, as the OECD average says 71% of men define their health status as "good" or "very good", however, this is only 67% for women. The differences between men and women are highest in countries such as Portugal, Turkey, France, United Kingdom. The answers are also changing according to age and social status. As it is expected older adults, also unemployed, having less education or income people state bad health status. In OECD countries, nearly 78% of adults,

with an available income in the top 20%, report their health as "good" or "very good". However, 61% of those, with an available income in the bottom 20%, give same answers.

As a result, it could be mentioned that self-reported health percentages obtained in each country could be explained by different factors. Each country must investigate the factors under people feelings about their health statuses. These feelings could be shaped on economical, sociological, physiological, even if climatic etc. factors. Each country must determine own policies for enhancing LS level in their home.

6. REFERENCES

- [1] <https://www.oecd.org/forum/issues/forum-issue-better-life-index.htm>.
- [2] OECD (2017), How's Life? 2017: Measuring Well-being, OECD Publishing, Paris. http://dx.doi.org/10.1787/how_life-2017-en
- [3] <https://www.oecd.org/social/the-path-to-happiness-lies-in-good-health-and-a-good-job-the-better-life-index-shows.htm>
- [4] <http://www.oecdbetterlifeindex.org/topics/life-satisfaction/>
- [5] E. Gundogan Aşık, A. Altın Yavuz, "Investigation of life satisfaction in OECD countries with multivariate analysis method", Journal of Social and Humanities Sciences Research (JSHSR), Vol. 5, no. 26, pp. 2547-2561, 2018.
- [6] L. Osberg and A. Sharpe, "An index of economic well-being for selected OECD countries", Review of Income and Wealth, vol. 48, pp. 291-316, 2012.
- [7] A. Kerényi, "The better life index of the organisation for economic co-operation and development", Public Finance Quarterly, vol. 56, pp. 518-538, 2011.
- [8] J. Kasparian and A. Rolland, "OECD's better life index: Can any country be well ranked?", Journal of Applied Statistics, vol. 39, pp. 2223-2230, 2012.
- [9] B. Stevenson and J. Wolfers, "Subjective Well-Being and Income: Is There Any Evidence of Satiation", American Economic Association, vol. 103, pp. 598-604, 2013.
- [10] H. Mizobuchi, "Measuring world better life frontier: a composite indicator for OECD better life index", Soc Indic Res, vol. 118, pp. 987-1007, 2014.
- [11] S. Akar, "Türkiye'de daha iyi yaşam endeksi: OECD ülkeleri ile karşılaştırma". Journal of Life Economics, vol. 1, no. 1, pp. 1-12, 2014.
- [12] M. Durand, "The OECD better life initiative: how's life and the measurement of well-being", Review of Income and Wealth, vol. 61, no. 1, pp.4-17, 2015.
- [13] O. Başol, "An evaluation on life satisfaction in OECD countries ", "IS, GUC" Industrial Relations and Human Resources Journal, vol. 20, no. 3, pp. 71-86, 2018.
- [14] E. Polat, "Determination of the effective economic and/or demographic indicators in classification of European Union member and candidate countries using Partial Least Squares Discriminant Analysis", Journal of Data Science, vol. 16, no. 1, pp. 79-92, 2018.
- [15] M. Barker, W.S. Rayens, "Partial least squares for discrimination", Journal of Chemometrics, vol. 17, pp. 166 - 173, 2003.
- [16] M.R. Almeida, D.N. Correa, W.F.C. Rocha, and F.J.O. Scafi, "Discrimination between authentic and counterfeit banknotes using Ramanspectroscopy and PLS-DA with uncertainty estimation", Microchemical Journal vol. 109, pp. 170-177, 2013.

- [17] H. Martens and M. Martens, "Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR)". *Food Quality and Preference*, vol. 11, no. 1-2, pp. 5-16, 2000.
- [18] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: a basic tool of chemometrics", *Chemometrics and Intelligent Laboratory Systems*, vol. 58, pp. 109-130, 2001.
- [19] B.M. Wise, N.B. Gallagher, R. Bro, J.M. Shaver, W. Windig, and R.S. Koch, R.S. "PLS Toolbox 4.0 for use with Matlab", 3905 West Eaglerock Drive, Wenatchee, WA, Eigenvector Research Inc. <http://www.eigenvector.com>, 2006.
- [20] D. Ballabio, F. Grisoni, and R. Todeschini, "Multivariate comparison of classification performance measures", *Chemometrics and Intelligent Laboratory Systems*, vol. 174, pp. 33-44, 2018.
- [21] D. Ballabio, R. Todeschini, "Chapter 4- Multivariate Classification for Qualitative Analysis" in Book: "Infrared Spectroscopy for Food Quality Analysis and Control" , pp.83-104, 2009. <https://doi.org/10.1016/B978-0-12-374136-3.00004-3>
- [22] Classification toolbox for MATLAB. file:///C:/Program%20Files/MATLAB/MATLAB%20Production%20Server/R2015a/classification_toolbox_5.2/help/index.htm
- [23] D. Ballabio and V. Consonni, "Classification Tools in Chemistry. Part 1: Linear Models. PLS-DA", *Analytical Methods*, vol. 5, pp. 3790-3798, 2013.
- [24] XLSTAT, Partial Least Squares Discriminant Analysis PLSDA Tutorial. https://help.xlstat.com/s/article/partial-least-squares-discriminant-analysis-plsda-tutorial?language=en_US