Representation of Socio-Economic Indicators for the Determination of Health Status

Volkan Kepoglu

Abstract

Socio-economic indicators can be used as health data, because there is a direct relationship and statistically correlation between socio-economic indicators and health status of the region. Some of the socio-economic indicators are more sensitive than the others for the determination of the health status in the region. The aim of this study is to find out which socio-economic indicators are more representative for the health status in Turkey. If the sensitive socio-economic indicators on health are defined for Turkey, the comparison of the health status between the provinces can be made. This empirical comparison will help the decision makers to locate and distribute health facilities in a more reasonable way.

The study will try to show the differentiation of the areas, which are underserved or over served on the health services. The indicators can be grouped into three subjects; demographical characteristics of the region, the rank for socio-economical development indexes and health information about the region. Statistical part of the study is based on these three subjects and spatial analysis is based on the administrative boundary of the province.

1. Introduction

The decision maker can politically take populist decisions on the health investments for the location of the new health facilities. However, decision maker in public sector, theoretically, needs to know overall the quality of health service per province. This situation has a priority for the rightful decision in which place the health investments should be allocated, but, there is no significant evidence for the determination of the total quality of the health service per each city. Moreover, to collect, manage and administrate the health data for the whole country with addition to the spatial component might be too difficult and expensive. In addition to this, health services are served separately by different health authorities in each province, which results to increase the complexity of the collection of health data.

In order to overcome these difficulties, socio-economic indicators can be used as health data, because there is a direct relationship and statistically correlation between socio-economic indicators and health status of the region. Some of the socio-economic indicators are more sensitive than the others for the determination of the health status in the region. The aim of this study is to find out which socio-economic indicators are more representative for the health status in Turkey. If the sensitive socio-economic indicators on health are defined for Turkey, the comparison of the health status between the provinces can be made. This empirical comparison will help the decision makers to locate and distribute health facilities in a more reasonable way.

As a result, the study will try to show the differentiation of the areas, which are underserved or over served by the health services. Under this objective specified above, the indicators can be grouped into three subjects; namely demographical characteristics of the region, the rank for socio-economical development indexes, and health information about the region. Statistical part of the study is based on these three subjects and spatial analysis is the administrative boundary of the province.

2. Study Area

The study area is the administrative boundary of the province that has 81 provinces in Turkey in 2004. The number of province in 1996 is 76. New provinces are formed in time from the different districts that are previously ties to different provinces. Therefore, if the statistical value in 1996 is the sum for the whole province and has no relation with the district, it could not be carried to 81 provinces in 2004. For this reason, Yalova, Düzce, Karabük, Osmaniye and Kilis provinces have no value in some analysis and coded as missing value.

3. Data Collection

In order to find out which socio-economic indicators are more sensitive for the representation of the health status in the region, many variables are collected and integrated to the GIS. The variables can be investigated under three logical subjects: demographical characteristics of the region, the rank for socio-economical development indexes and health information about the region.

All collected variables have as a spatial component. Some spatial components are based on the province and some are on hospital. Hospital based information has aggregated as a summation to the province. All variables are structured in the form of table and integrated into the personal geodatabase, which is the MS Access MDB format and known as the GIS format of ESRI's product, so, all feature layers and tables are stored in one file. Structured tables are also used for statistical analysis.

The feature layers listed below are used during the analyses. The definition about them is that the feature name shows the name of the feature layer in the personal geo-database and source of the data explains from where the statistical data has been taken. Features are grouped according to their related subjects as follows;

A. List of Feature Layers on the Demographical Characteristics of the Province

1. Infant mortality rate in 2000 (0-1 age)

   Feature Name: INFANTDEATHRATE2000

   Source of Data: General Directorate of Regional Development and Structural Adaptation in State Planning Organization
2. The number of retired and insured person in 2000 from three different social security organizations

   Feature Name: RETIREDINSURED2000

   Source of Data: General Directorate of Regional Development and Structural Adaptation in State Planning Organization
3. Mortality rate with distribution of age groups in 1999

   Feature Name: PERSONMORTALAGE1999

   Source of Data: Socio-economical Indicators in State Institute of Statistics
4. Total, rural, urban population and annual population growth rate in 1990 and in 2000

   Feature Name: POPULATION2000

   Source of Data: General Directorate of Regional Development and Structural Adaptation in State Planning Organization

B. List of Feature Layers on the Effect of Socio-economical Development in Province

1. The rank for socio-economical development in 1996

   Feature Name: DEVELOPMENT1996

   Source of Data: The Work for Socio-economical Development in State Planning Organization
2. The rank for socio-economical development in 2003

   Feature Name: DEVELOPMENT2003

   Source of Data: General Directorate of Regional Development and Structural Adaptation in State Planning Organization

   Two socio-economical development feature layers far different years are the result of 58 different variables. The variables range from fertility rate, electricity consumption per people to number of car to fax per person. All these variables were used in the socio-economical development analysis that is carried out by the State Planning Organization. As a result of that work, development index has been calculated for each city. In this analysis, this development index has been accepted as an input, because the index is the representation of all these variables.

3. Governmental investment in health sector in 2001

   Feature Name: INVESTMENT2001

   Source of Data: Socio-economical Indicators in State Institute of Statistics

C. List of Feature Layers on Health Information for the Province

1. The number of beds and hospitals in public and private sector in 2000

   Feature Name: BEDHOSPTNUM2000

   Source of Data: Socio-economical Indicators in State Institute of Statistics
2. Treatment statistics of the Social Security Institutions , in 2001 and in 2003 (number of inpatients, days in bed, number of people dismissed, number of deaths, number of applications for policlinics, number of examinations in policlinics, number of prescriptions, number of medicine, medicine per prescription, rate of bed occupation, average time of using beds, interval bed cycle, hospital crude death rate, unused days in bed, days in bed for death and dismissed people, days in bed necessary to use)

   Feature Name: SSKCURE2001 and SSKCURE2003

   Source of Data: Social Securities Institution
3. Characteristics of all hospitals in Turkey in 2003 (bed capacity, bed capacity for intensive clinic, number of operating room, daily number of clinic for out-patients, daily number of surgical operation, available type of radiological equipment, availability of specialized centers, type of hospital, number of hospital in province)

   Feature Name: HOSPTPUBLIC2003

   Source of Data: Hospital Guide, 2003
4. The number of health personnel in 1999 and in 2000 (physician, dentist, nurse, midwife, health officer, pharmacist)

   Feature Name: HLTHIND1999

   Source of Data: Socio-economical Indicators in State Institute of Statistics

   Feature Name: HLTHIND2000

   Source of Data: General Directorate of Regional Development and Structural Adaptation in State Planning Organization

4. Data Analysis

All variables, listed above, should be integrated into the GIS layer in order to start the analysis. This can be done by transforming the structure of data from table based format to GIS based format. At the end, all variables should be seen as fields in the attribute table of feature layer. For the GIS software, one of the most powerful products of ESRI, ArcInfo 9.0 has been chosen. ArcMap and ArcCatalog are the software that is part of this ArcInfo package.

At first, all variables are structured in the form of database table by using MS Excel. For the GIS layer, a unique value has assigned to the each province. And then, the same values for all tables are written for the each province in MS Excel. By ArcCatalog, a personnel geodatabase has been created. With the help of MS Access, all excel tables are imported into the personnel database. Each table is registered as Geodatabase by using ArcCatalog. The table registration gives the access and editing permissions to the ArcMap. Using this common field that exists on both layers, relation between the feature layer and table has been made in ArcMap. After the connections between them, all features are export to SHP files. This makes the tables a direct relation inside the native format of GIS layer. Finally, all SHP files are imported to personnel geodatabase by using ArcCatalog. By doing all these transformations, data become ready to use with the GIS and display the output as a map.

All feature layers are classified into two groups; demand and supply part of the health service provision according to the meaning of the variables. By this way, they will become the representative layer either on the demand side or on the supply side. By the help of factor analysis, it will be found out that how much relation each variable has on the representation of the demand or supply of health services.

The number of health personnel in 1999 and in 2000, the properties of all hospitals in Turkey in 2003, the number of bed and hospital in public sector in 2000, governmental investment in health sector in 2001, the treatment statistics of SSK in 2001 and in 2003 are related on the supply side. Infant mortality rate in 2000, mortality rate with the distribution of age groups in 1999, socio-economical development rank in 1996 and in 2003, the number of retired and insured person in 2000 from three different social security organizations and total, rural, urban

population and annual population growth rate in 1990 and in 2000 are related on the demand side of the analysis. From the grouped values, two summary tables have been prepared. One of them is for the use of health service, and the other one is for the provision of health service. These summary tables are exported to the statistical program for the application of factor analysis. Factor analysis will give the importance of variables in degree by specifying the relationship between them.

Standardization should be applied to each variable, because each value has different units. By doing the standardization, different weights among the variables will be avoided. In the summary tables, mean and standard deviation have been calculated separately for each field. Mean has been subtracted from the each value of the field and divided by the standard deviation. This process has been done both for the demand and supply tables.

When the demand summary table is analyzed by the factor analysis with the statistical program, within the range of %86.32 four components is enough to explain the relationship as shown below. The composition of the each component and the rate of their relation are also displayed below. Negative value shows that there is a reverse relationship between the variable and the component.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.667	48.150	48.150	8.667	48.150	48.150	6.107	33.925	33.925
2	3.524	19.579	67.729	3.524	19.579	67.729	4.196	23.313	57.238
3	2.061	11.450	79.179	2.061	11.450	79.179	2.812	15.619	72.857
4	1.286	7.145	86.324	1.286	7.145	86.324	2.424	13.467	86.324
5	.791	4.392	90.716						
6	.603	3.348	94.063						
7	.480	2.666	96.730						
8	.263	1.463	98.193						
9	.167	.925	99.118						
10	.062	.346	99.464						
11	.049	.272	99.736						
12	.018	.102	99.839						
13	.011	.064	99.902						
14	.009	.052	99.954						
15	.005	.027	99.981						
16	.003	.017	99.998						
17	.000	.002	100.000						
18	-6,73E-16	-3.740E-15	100.000						

Extraction Method: Principal Component Analysis.

Rotated Component Matrix[a]

	Component			
	1	2	3	4
sSEHR2000	.967			
sTPLM2000	.962			
sSSKTPL01	.954			
SOLYAS1	.945		.241	
SOL_TPL	.930	.206	.212	
SENDEKS	.736	.620		
SKOY2000	.596			.470
SGELIS	-.254	-.890		
SGELIS03	-.319	-.844		
SBBKÖLÜM		-.779		
sYS65US_N	.266	.751	.373	-.242
SYS0_14N			.933	
sYS15_64N		.556	.742	
SOL_NFS		.638	.724	
SYS0_N	.387		.676	
sTOPLAMHI	.208			.938
sSEHIRHIZ		-.425		.762
SKOYHIZ	.453	.267	.208	.732

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

As it can be seen from rotated component matrix, the context of 1st ordered component in the demand side is mainly composed of urban population, total population, total number of retired and insured person in the SSK, total mortality rate per person. These variables are respectively ordered from strong to weak relationship in the first order. Within the similar meaning, second ordered component is socio-economical development in 1996 and in 2003, and mortality rate over 65 ages. Socio-economical rank has negative values, which possibly means that the most developed cities have relatively lower demand to the health services.

The same thing is true for the supply part. Seven components are required to explain the relationship between the component and the variables within the range of %86.231. The results are displayed below.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	20.306	48.346	48.346	20.306	48.346	48.346	18.839	44.854	44.854
2	6.092	14.504	62.850	6.092	14.504	62.850	6.732	16.028	60.881
3	3.963	9.436	72.286	3.963	9.436	72.286	4.242	10.099	70.981
4	2.563	6.101	78.387	2.563	6.101	78.387	2.785	6.632	77.613
5	1.200	2.856	81.243	1.200	2.856	81.243	1.345	3.201	80.814
6	1.059	2.521	83.764	1.059	2.521	83.764	1.201	2.859	83.673
7	1.036	2.467	86.231	1.036	2.467	86.231	1.075	2.558	86.231
8	.925	2.202	88.433						
9	.775	1.846	90.279						
10	.667	1.588	91.867						
11	.613	1.460	93.327						
12	.473	1.126	94.454						
13	.366	.872	95.326						

Rotated Component Matrix[a]

	Component						
	1	2	3	4	5	6	7
Zscore(KM_YTK00)	.983						
Zscore: Sum_GAMLYT	.979						
Zscore(ECZANE99)	.977						
Zscore: YATTPLM_NFS01	.973						
Zscore: Sum_PEOPLE	.961	.207					
Zscore: Sum_INPATI	.961	.210					
Zscore(SUM_BEDS)	.958						
Zscore: MRKZVAR_SA	.951						
Zscore: Sum_GPLKLN	.948						
Zscore: Sum_AP_PCL	.948	.258					
Zscore: ALLIHTISAS	.941						
Zscore: Sum_PRESCR	.931	.267					
Zscore: Sum_MEDICI	.928	.258					
Zscore: Sum_DEATHS	.921					-.218	
Zscore: Sum_YOGUNB	.919			.211			
Zscore: RCHZ_TYPSA	.918						
Zscore: Sum_YBAKIM	.884			.237		.220	
Zscore: HSTN_ADET	.877	.295					
Zscore: KM_HAST00	.866					.270	
Zscore: Sum_YENID	.856						
Zscore: Sum_KUVOZ_	.834			.290		.280	
Zscore: MEDICI_NFS		.948					
Zscore: PRESCR_NFS		.940					
Zscore: AP_PCL_NFS		.933					
Zscore: INPATI_NFS	.216	.923					
Zscore: PEOPLE_NFS	.219	.922					
Zscore: DEATH_NFS	.381	.750					
Zscore: Ave_R_BED_	.240	.690		.218		.261	-.255
Zscore: Ave_CRD_DE	.207	.528					-.366
Zscore: HMSR_NFS00			.919				
Zscore: HKM_NFS00			.886				
Zscore: EBE_NFS00			.850	.212			
Zscore: SGLKC_NFS00			.820			-.267	
Zscore: DSHKM_NFS00			.775			.215	
Zscore: ECZCI_NFS00			.686			.300	
Zscore: YATSGLK_NFS01	.217	.203		.814			
Zscore: YATYZD_NFS01				.808			
Zscore: GAMLYSAY_N	.442	.272		.631	-.277		
Zscore: GPNKSAY_NF		.377		.583	-.350		
Zscore(BEDS_NFS)					.847		
Zscore: KMYTK_NF00		-.430			.479	-.501	
Zscore: Ave_BEDCYC		.258		-.299			.698

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

For the supply part, the values in the rotated component matrix are so close. This shows that chosen variables have a strong relationship between each other, and they will be used to determine the relationship of the supply provision for the health services. In the first ordered component, bed capacity of the public sector, daily number of surgical operation per person, the number of private drugstore, total governmental investment in the health sector, total number of people dismissed from the hospital, number of inpatient, availability of specialized centers, daily number of clinic for out-patients, number of applications for policlinics, availability of specialized centers, number of prescriptions per person, number of medicine per person, mortality rate, bed capacity for intensive clinic, available type of radiological

equipment, total number of hospital per province, number of public hospital are the variables that have the closest and highest values for the representation of the supply of health services.

5. Discussion

In this study, choosing which variables should be used is the most critical and sensitive decision. Selection of variables has been done according to the post graduate study that the similar methodology has been applied for the Canada. Some of the variables in the following subjects such as the coverage and quality of disease prevention, reproductive health and nutrition, are the widely accepted indicators for the health status in the world. However, this study has become limited with the available data, because for some of the variables, it could not be reached or may not exist as statistical information among the public institutions.

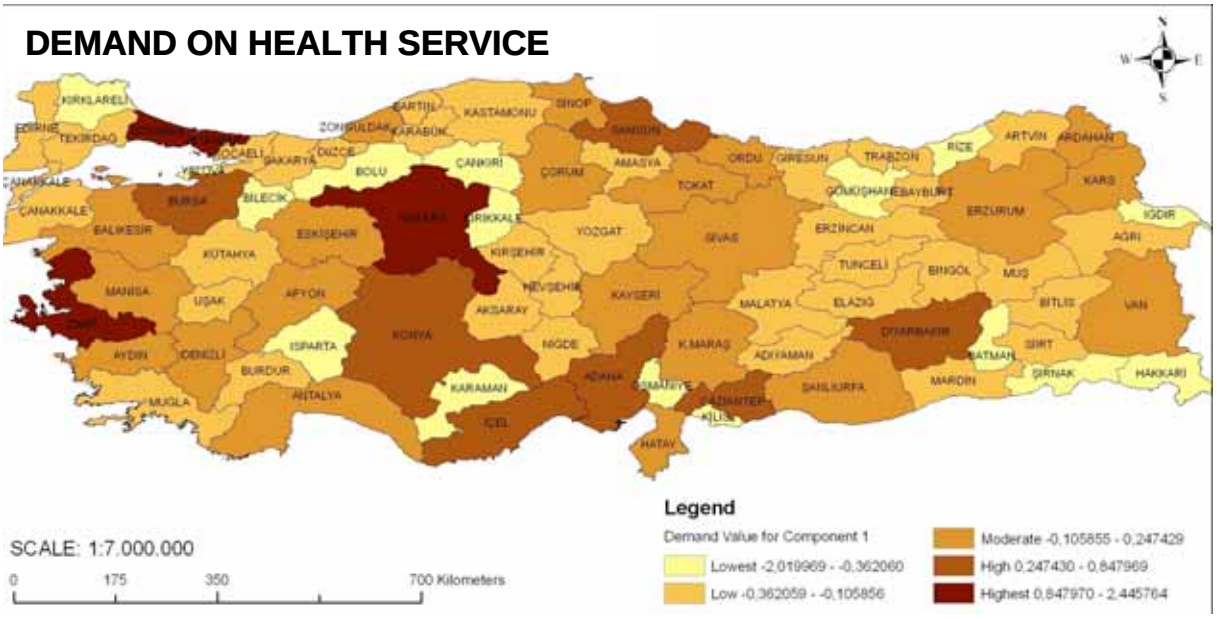
The results and the selection of the variables have only become valid for this study and specific for the country. Within these considerations, it should be preferred to ask more health officer's expertise for the appropriate selection of the variables. Especially for the demand part, it will be thought that there may not be enough variables. However, based on the same methodology with the increased number of variables, the study can be enriched very easily.

As a result, the analysis serves for the preliminary study and for the use of the methodology. It could be said for the study that it is an appropriate example for the integration of statistical variables into the GIS environment and the implementation of the methodology for finding out the representation of the variables according to the results of factor analysis. The degree of the representation for the chosen socio-economical variables should be tested with the geo-coded health data. This comparison can be the subject of another study.

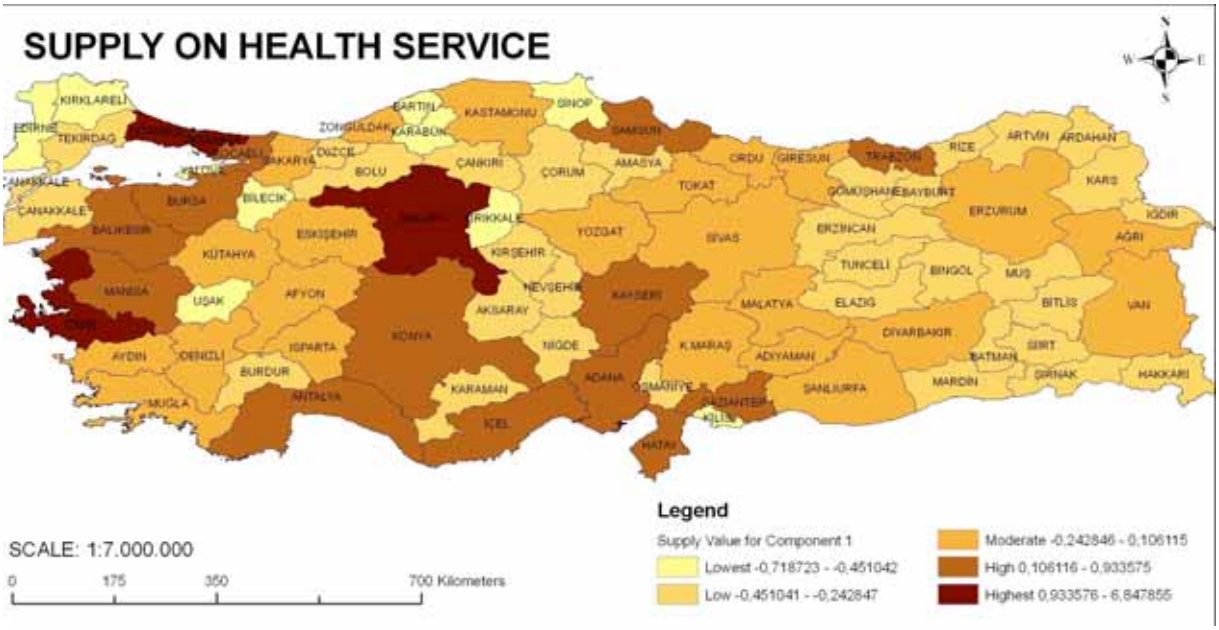
6. Results and Conclusion

Two figures show the mapping result of first ordered component as seen below. First figure is for the demand side in health service and the last one is for supply side.

The most developed cities, Istanbul, Ankara and Izmir, respectively, have the highest demand value on the health service. The second demand group is composed from the provinces of Bursa, Samsun, Konya, Icel, Adana, Gaziantep and Diyarbakir. These cities are the second ordered developed cities. Third group is Balikesir, Manisa, Aydin, Denizli, Afyon, Eskisehir, Zonguldak and Antalya. Although Antalya is the most developed cities in tourism, the city needs the health service in the third order when it is compared with the other cities.



On the supply side, the same cities are valid in the first group, which are Istanbul, Izmir and Ankara. It means that these cities provides high amount of health services, and also they requires more demand. The second group is composed from the provinces of Bursa, Balikesir, Manisa, Konya, Antalya, Icel, Adana, Kayseri, Hatay, Gaziantep, Samsun and Trabzon.



When the comparison between the demand and the supply has been made, it can be seen that Trabzon and Kocaeli provinces are the over served areas for the provision of health service. Many cities are in the same group, however, the most delicate provinces are Edirne, Bartin, Karabük, Usak, Corum, Sinop, Diyarbakir, Ardahan and Kars. These provinces are the under-served areas for the provision of health service. Sinop is the most under-served provinces for the provision of health service.

7. References

1. SPSS Base 9.0 Applications Guide

2. GIS Based Statistical Health Risk Model for Calgary Communities, Woods, Teresa Marie, the Degree of Master of Science, Department of Geography, Calgary, Alberta, December 1997

3. Planning for Health Care: A New Orleans, Louisiana Case of Spatial Data and Methods to Increase Accessibility and Accountability in the Renal Care Industry, Neu, Denese Maureen, the Doctor of Philosophy, College of Urban and Public Affairs, December 2002

4. Hospital Guide, 2003

5. Treatment Statistics, Social Securities Institution

6. Socio-economical Indicators, State Institute of Statistics

7. The Work for Socio-economical Development, 1996, State Planning Organization

8. General Directorate of Regional Development and Structural Adaptation, State Planning Organization

9. World Development Indexes, 2003, World Bank

About Author

Volkan Kepoglu

Title: Mapping and GIS Manager, in Baku-Tbilisi-Ceyhan Crude Oil Pipeline Project and Post Graduate Student, in the Department of Geodesy and Geographic Information Technologies, Middle East Technical University

Address: BOTAS Pipeline Corporation, Bilkent Plaza, A1 Block, 06530, Bilkent, Ankara, Turkey.

Work Tel: +90 312 297 31 66

Cellular Phone: +90 532 705 65 71

Fax: +90 312 266 06 10

E-mail: volkan.kepoglu@botas.gov.tr

vkepoglu@yahoo.com