Scholarly Research Journal for Interdisciplinary Studies,
Online ISSN 2278-8808, SJIF 2016 = 6.177, www.srjis.com
UGC Approved Sr. No.49366, SEPT-OCT 2017, VOL- 4/36

GEOSPATIAL GROUPING ANALYSIS OF SEVERITY OF NEEDS AT DISTRICT LEVEL IN ETHIOPIA

Neerja Asthana

Associate Professor, Department of Military Studies, Bareilly College, Bareilly. India. Email: neerajaasthana37@gmail.com

Abstract

The objective of this study was to study spatial variation in severity of need at zone level in Ethiopia. To meet the objective, grouping of the zones to find natural clusters in case of the severity of the need was performed. Four variables were chosen under severity of need as nutrition, food, agriculture and the health. Ethiopian government approved zone boundaries were used in this study. The data was collected from the Data Repository of the United Nation's OCHA. To make the five groups of the zones based on severity of the need, grouping analysis tool of ArcGIS platform was used. Grouping effectiveness was measured using the Calinski-Harabasz pseudo F-statistic. The R2 values obtained for nutrition (0.86), food (0.85), agriculture (0.78) and the health (0.65). The larger obtained the R2 value shows that variables are at discriminating among the zones. Overall distribution is positively skewed, because the whisker and half-box are longer on the right side of the median than on the left side. Group wise results shows the 55 zones under group two are in satisfactory mode however another groups are in a severe mode of the need. Variable wise results indicate nutrition need is most severe in group 6 and group 3. Food need is severe in group 1, 3, 4 & 5. Agriculture need is severe in group 1, 3 & 5. Health needs are more severe in group 1, 3 &5. The research work will serve the actors engaged in humanitarian aid to reduce the poverty in Ethiopia.

***Keywords**: OSCHA, F-statistic, Grouping Analysis, Geostatistical analysis.*

1. Introduction

The drought conditions in Ethiopia are very frequent. Ethiopia has experienced its second severe drought in less than two years. Insufficient rainfall during the 2017 rainy season has led to severe water shortages, catastrophic livestock losses, and failed crops throughout the country .

In addition to drought, populations across Ethiopia face other challenges that contribute to sustained humanitarian needs and an ongoing complex emergency including above-average food prices, disease outbreaks, localized intercommunal conflict, seasonal flooding, and limited access to health and water, sanitation, and hygiene services.

By March 2018, the Government of Ethiopia and UN estimated that nearly 7.9 million people required emergency food assistance, including 1.8 million people in the country's Somali Region. Drought has displaced thousands of people in Somali, impeding access to basic services (USAID, 2018).

There are many tools available in ArcGIS for cluster analysis include Spatially Constrained Multivariate Clustering, Group analysis, Multivariate Clustering, Density-Based Clustering, Image Segmentation, Hot Spot Analysis, Cluster and Outlier Analysis tools, and the Space Time Pattern Mining tools (Bennett, 2018). In this work group analysis tool was used.

In this paper a new approach was adopted to investigate spatial variation in need analysis. The present paper aims to make the spatial groups based on the nature of the need. The aim of our work was also to broaden current knowledge of need in Ethiopia.

The paper calls into sheds new light on this issue. With this in mind we undertook this study and developed grouping analysis methodology to study need problem in Ethiopia. The findings of the work will help the policy makers to develop the programmes to minimise the food problem and to minimise the deaths rate in hard hit zones by providing infrastructure in the food storage centres.

2. Materials and methods

2.1 Study Area

The Ethiopia was chosen as the study area. This is the second largest populous country of the Africa, covering about 0.8 million square kilometre of the area. There are ten regional states and two chartered cities, the latter being the country's capital Addis Ababa, and Dire Dawa, which was chartered in 2004. Being based on ethnicity and language, rather than physical geography or history, the regions vary enormously in area and population; the most notable example is the Harari Region, which has a smaller area and population than either of the chartered cities. When they were established in 1992, there were more regions, but five regions were merged to form the multi-ethnic Southern Nations, Nationalities, and Peoples' Region later in 1992, following the first elections of regional councils on 21 June 1992 (Lyons, 1996).

2.2 Data used

2.2.1 Countries Boundaries

Ethiopia government-approved administrative boundaries were used in this study.

2.2.2 Severity of the need data

The Severity of the need data was collected from the website https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx/resource/5837b491-f34a-40a7-8bb0-135330aa88da?view_id=d5bbcd0b-1867-4739-a93e-2b0fe886c6f4#. This data contains the severity of needs analysis prepared for the 2017 humanitarian needs overview. The severity score indicates how compounded the humanitarian needs in a woreda are. The higher the severity score of a woreda, the more severe, time-critical and compounded the needs are (OCHA, 2017).

Figure 2 shows the flow chart; first we converted the CSV format of data into shape files. Subsequently we analysed cases and formed the groups. Finally, Ordinary Least Squares (OLS) linear regression was performed.

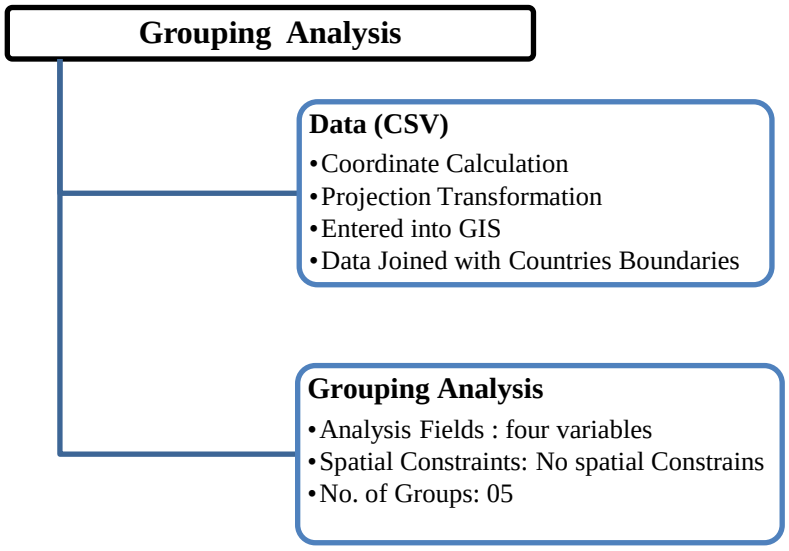

Fig. 2 the flowchart of the used methodology

2.3 Data Analysis

2.3.1 Grouping Analysis

Grouping Analysis performs a classification procedure that tries to find natural clusters in your data. Given the number of groups to create, it will look for a solution where all the features within each group are as similar as possible, and all the groups themselves are as different as possible. Feature similarity is based on the set of attributes that you specify for the Analysis Fields parameter and may optionally incorporate spatial properties or space-time properties. When space or space-time Spatial Constraints is specified, the algorithm employs a connectivity graph (minimum spanning tree) to find natural groupings. When

NO_SPATIAL_CONSTRAINT is specified, the Grouping Analysis tool uses a K Means algorithm.

Grouping effectiveness is measured using the Calinski-Harabasz pseudo F-statistic, which is a ratio reflecting within-group similarity and between-group difference as follows:

$$\text{Calinski} - \text{Harabasz pseudo F} - \text{statistic} = (R^2/n_c - 1)/(1 - R^2/n - n_c) \quad (1)$$

Where:

$$R^2 = \frac{SST - SSE}{\text{SST}} \quad (2)$$

and SST is a reflection of between-group differences and SSE reflects within-group similarity:

$$\text{SST} = \sum_{I=1}^{nc} \sum_{k=1}^{ni} \sum_{j=1}^{nv} (V_{ij}^k - \overline{V^k})^2 \text{.....3}$$

$$\text{SST} = \sum_{I=1}^{nc} \sum_{k=1}^{ni} \sum_{j=1}^{nv} (V_{ij}^k - \overline{V_i^k})^2 \text{.....4}$$

n = the number of features

n_i = the number of features in group t

n_c = the number of classes (groups)

n_v = the number of variables used to group features

V_{ij}^k = the value of the k^{th} variable of the j^{th} feature in the t^{th} group

$\overline{V^k}$ = the mean value of the k^{th} variable

$\overline{V_i^k}$ = the mean value of the kth variable in group i

3. Results

3.1. Group wise results

The distribution is positively skewed, because the whisker and half-box are longer on the right side of the median than on the left side. The first set of summary statistics are printed in black because these are the global Mean, Standard Deviation (Std.Dev.), Minimum, Maximum, and R2 values for all data in each analysis field. The larger the R2 value is for a particular variable, the better that variable is at discriminating among features. After the global summaries, the Mean, Standard Deviation, Minimum, Maximum, and Share values are reported for each variable in each group.

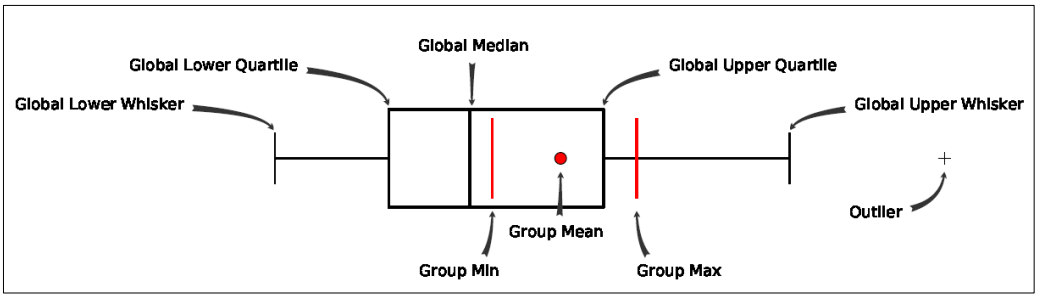

Fig. 3 Overall Statistics

The results of Grouping Analysis are summarised in table 1.

Table 01 Overall Variable Statistics: Std. Distance = 0.2576, SSD = 76.9940

Variable	Mean	Std. Dev.	Min	Max	R2
NUTRITION	0.0748	0.0546	0.0458	0.35	0.8665
FOOD	0.2086	0.1607	0.05	0.7276	0.8543
AGRICULTURE	0.2061	0.1814	0.05	0.9	0.7873
HEALTH	0.1038	0.0682	0.025	0.4031	0.655

Data Source: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx

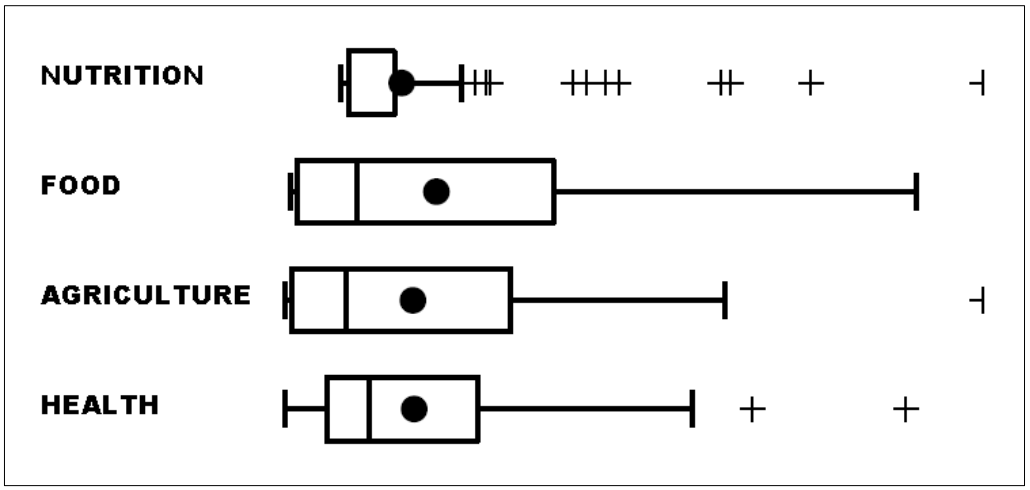

Fig. 4 Overall Variable Statistics

Group wise results

Group one (blue) zones are shown in blue colour, these zones are Afder, Awsi /Zone 1, Bale, Doolo, Erer, Fafan, Fanti /Zone 4, Gabi /Zone 3, Guji, Jarar, Kemashi, Kilbati /Zone2, Korahe, Nogob, Siti, West Guji. These zones have larger need of food.

Table 2 Group 1: Count = 16, Std. Distance = 0.1538, SSD = 20.3773

Variable	Mean	Std. Dev.	Min	Max	R2
NUTRITION	0.0796	0.0217	0.0458	0.1167	0.2329

FOOD	0.3902	0.0855	0.29	0.5805	0.4288
AGRICULTURE	0.3549	0.1183	0.1917	0.5852	0.463
HEALTH	0.1683	0.0435	0.1022	0.2727	0.451

Data Source: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx

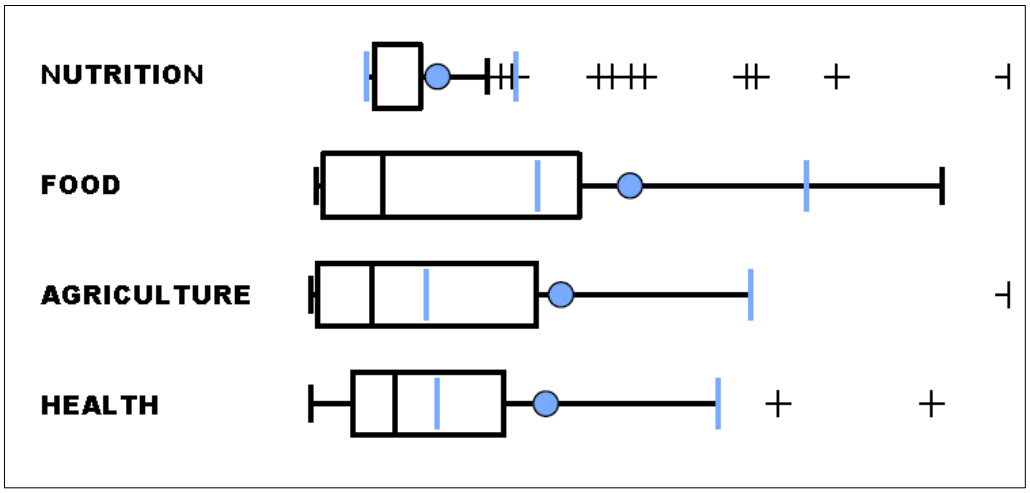

Fig. 5 Group 01 Statistics

Group two (red) is the far largest one comparing with another group. This group includes 55 zones out 92 zones of the Ethiopia, so it is matching with median values. This group include Agnewak, Alle, Arsi, Asosa, Awi, Basketo, Bench Sheko, Buno Bedele, Burji, Central, Central Gondar, East Gojam, East Shewa, East Wellega, Eastern, Finfine Special, Guraghe, Hadiya, Harari, Hari /Zone 5, Horo Gudru Wellega, Ilu Aba Bora, Itang Special woreda, Jimma, Kefa, Kelem Wellega, Konso, Majang, Mao Komo Special, Mekelle, Metekel, Mirab Omo, North Gondar, North Shewa (AM), North Shewa (OR), North Wello, North Western, Nuwer, Oromia, Region 14, Sheka, Siltie, South Eastern, South Gondar, South Omo, South Wello, South West Shewa, Southern, Wag Hamra, West Gojam, West Gondar, West Shewa, West Wellega, Western, Yem Special.

Table 3 Group two Count = 55, Std. Distance = 0.0814, SSD = 33.9389

Variable	Mean	Std. Dev.	Min	Max	R2
NUTRITION	0.0559	0.0169	0.0467	0.1625	0.3808
FOOD	0.0907	0.0566	0.05	0.2735	0.3299
AGRICULTURE	0.0821	0.0389	0.05	0.1967	0.1725
HEALTH	0.0784	0.0404	0.025	0.2111	0.4922

Data Source: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx

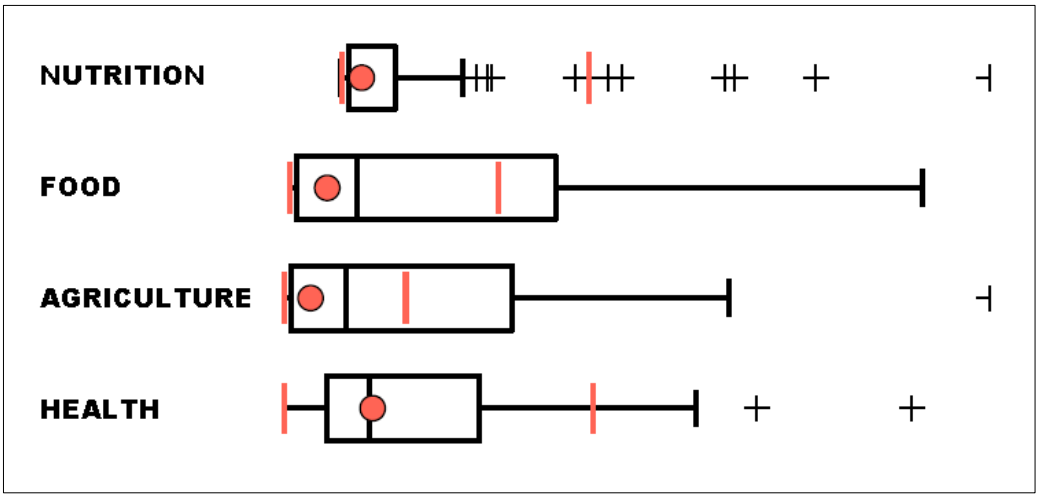

Fig. 6 Group 02 Statistics

Group three (green) include 07 zones. East Bale, East Hararge, Gedeo, Liban, Shabelle, West Arsi and West Hararge.

Table 4 Group three Count = 7, Std. Distance = 0.1474, SSD = 13.2939

Variable	Mean	Std. Dev.	Min	Max	R2
NUTRITION	0.2141	0.0429	0.1561	0.2686	0.3701
FOOD	0.3698	0.073	0.2173	0.4462	0.3377
AGRICULTURE	0.4	0.1056	0.1873	0.5201	0.3916
HEALTH	0.1977	0.0585	0.1108	0.3097	0.5259

Data Source: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx

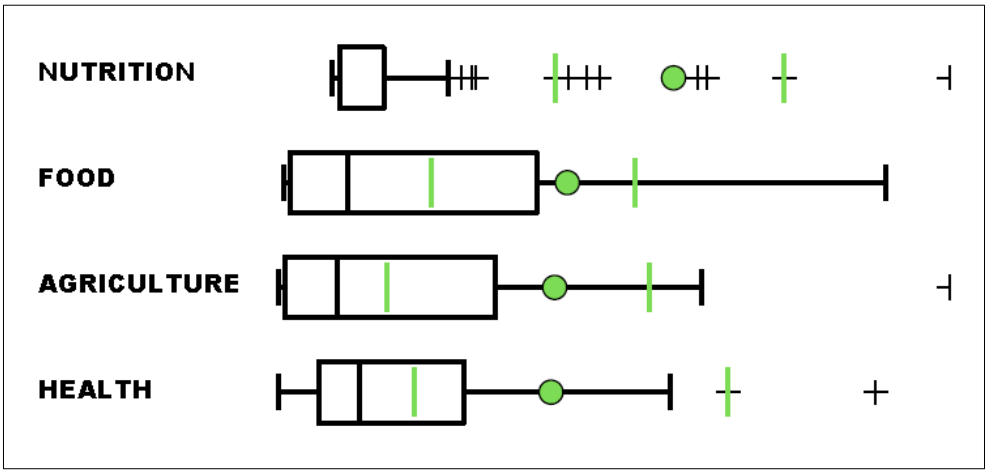

Fig. 7 Group 03 Statistics

Group four (orange) include 13 zones. These are Amaro, Borena, Dawuro, Derashe, Dire Dawa rural, Dire Dawa urban, Gamo, Gofa, Halaba, Kembata Tibaro, Konta Special, Sidama and Wolayita..

Table 5 Group four Count = 13, Std. Distance = 0.1464, SSD = 9.3840

Variable	Mean	Std. Dev.	Min	Max	R2
NUTRITION	0.0532	0.0058	0.0458	0.0708	0.0822
FOOD	0.3571	0.0351	0.3117	0.4068	0.1403
AGRICULTURE	0.3897	0.1411	0.1917	0.5726	0.4482
HEALTH	0.0584	0.0164	0.0447	0.103	0.154

Data Source: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx

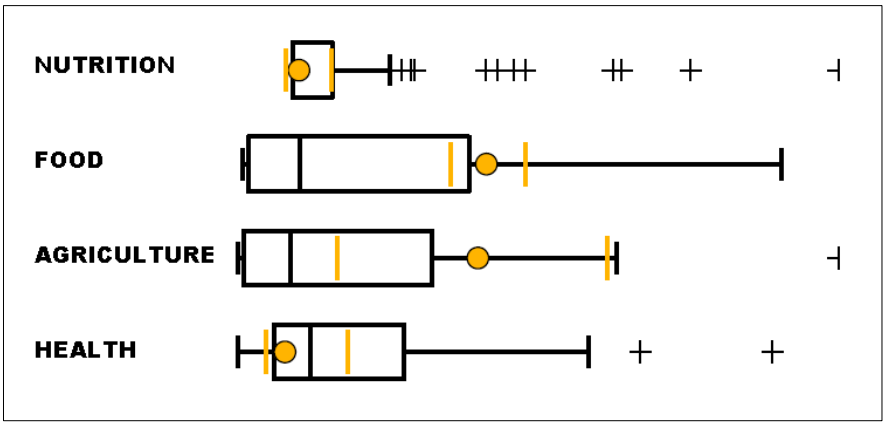

Fig. 8 Group 04 Statistics

Group five (Magenta) include only one zone Daawa.

Table 6 Group five Count = 1, Std. Distance = 0.0000, SSD = 0.0000

Variable	Mean	Std. Dev.	Min	Max	R2
NUTRITION	0.35	0	0.35	0.35	0
FOOD	0.7276	0	0.7276	0.7276	0
AGRICULTURE	0.9	0	0.9	0.9	0
HEALTH	0.4031	0	0.4031	0.4031	0

Data Source: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx

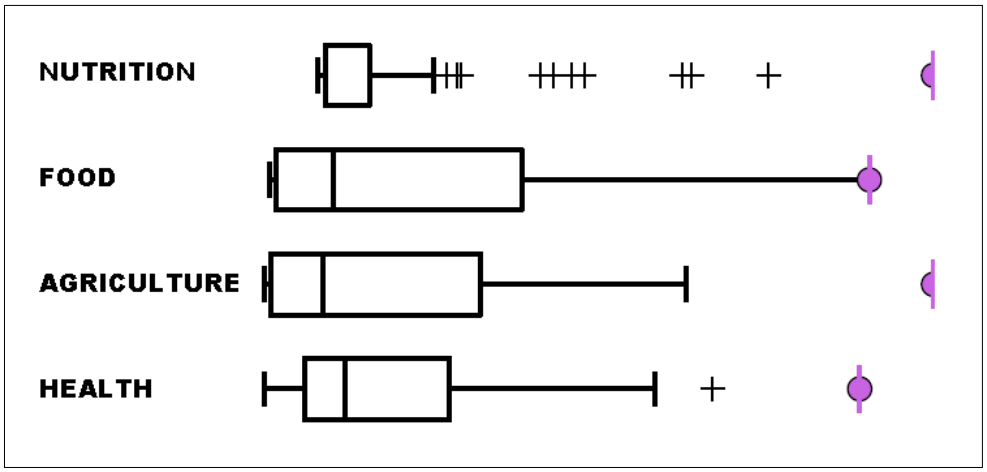

Fig. 9 Group 05 Statistics

3.2 Variable-Wise Results

Table 07 shows the result of the Nutrition variable.

Table 7 NUTRITION R2 = 0.87

Group	Mean	Std. Dev.	Min	Max	Share
1	0.0796	0.0217	0.0458	0.1167	0.2329
2	0.0559	0.0169	0.0467	0.1625	0.3808
3	0.2141	0.0429	0.1561	0.2686	0.3701
4	0.0532	0.0058	0.0458	0.0708	0.0822
5	0.35	0	0.35	0.35	0
Total	0.0748	0.0546	0.0458	0.35	1

Data Source: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx

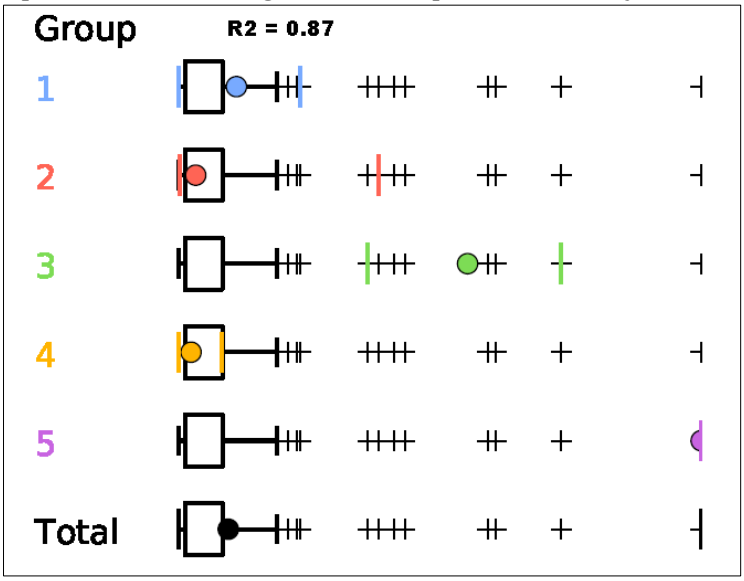

Fig. 10 Nutrition Statistics

Table 08 shows the result of the Food Variable

Table 8
FOOD : R2 = 0.85

Group	Mean	Std. Dev.	Min	Max	Share
1	0.3902	0.0855	0.29	0.5805	0.4288
2	0.0907	0.0566	0.05	0.2735	0.3299
3	0.3698	0.073	0.2173	0.4462	0.3377
4	0.3571	0.0351	0.3117	0.4068	0.1403
5	0.7276	0	0.7276	0.7276	0
Total	0.2086	0.1607	0.05	0.7276	1

Data Source: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx

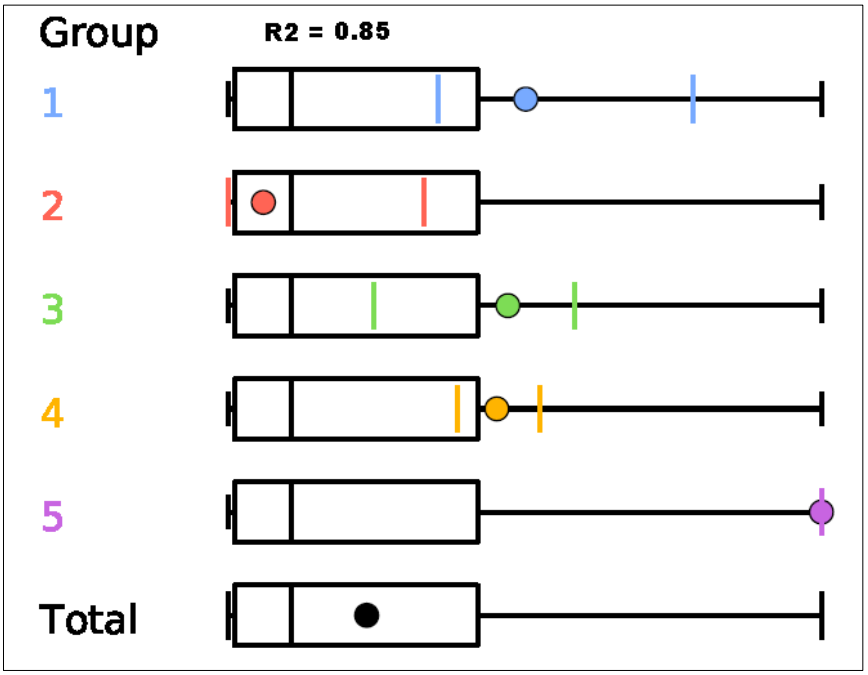

Fig. 11 Food needs Statistics

Table 09 shows the result of the Agriculture Variable

Table 9 AGRICULTURE : R2 = 0.79

Group	Mean	Std. Dev.	Min	Max	Share
1	0.3549	0.1183	0.1917	0.5852	0.463
2	0.0821	0.0389	0.05	0.1967	0.1725
3	0.4	0.1056	0.1873	0.5201	0.3916
4	0.3897	0.1411	0.1917	0.5726	0.4482
5	0.9	0	0.9	0.9	0
Total	0.2061	0.1814	0.05	0.9	1

Data Source: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx

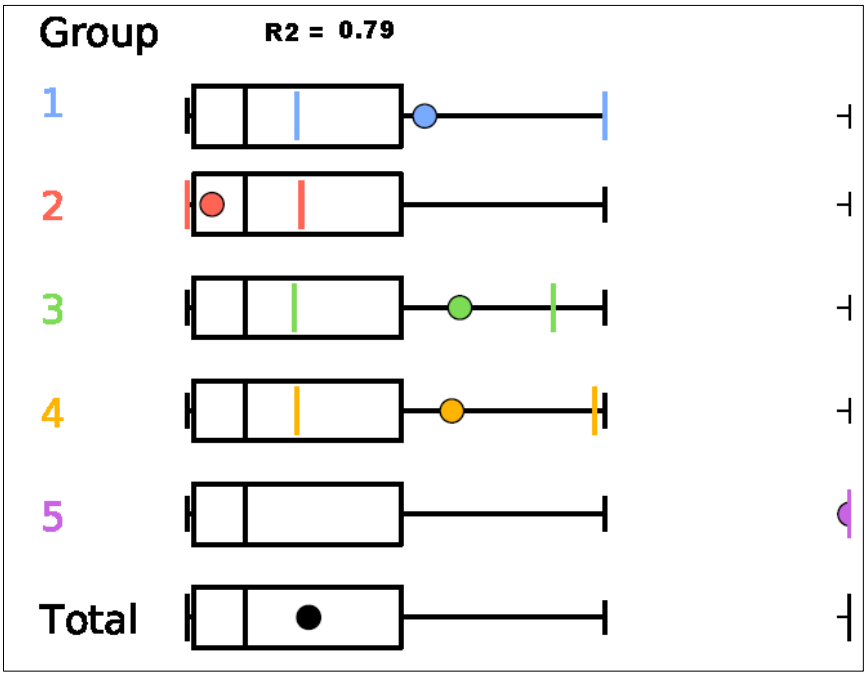

Fig. 12 Agriculture Needs Statistics

Table 10 shows the result of the health Variable

Table 10 HEALTH : R2 = 0.66

Group	Mean	Std. Dev.	Min	Max	Share
1	0.1683	0.0435	0.1022	0.2727	0.451
2	0.0784	0.0404	0.025	0.2111	0.4922
3	0.1977	0.0585	0.1108	0.3097	0.5259
4	0.0584	0.0164	0.0447	0.103	0.154
5	0.4031	0	0.4031	0.4031	0
Total	0.1038	0.0682	0.025	0.4031	1

Data Source: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx

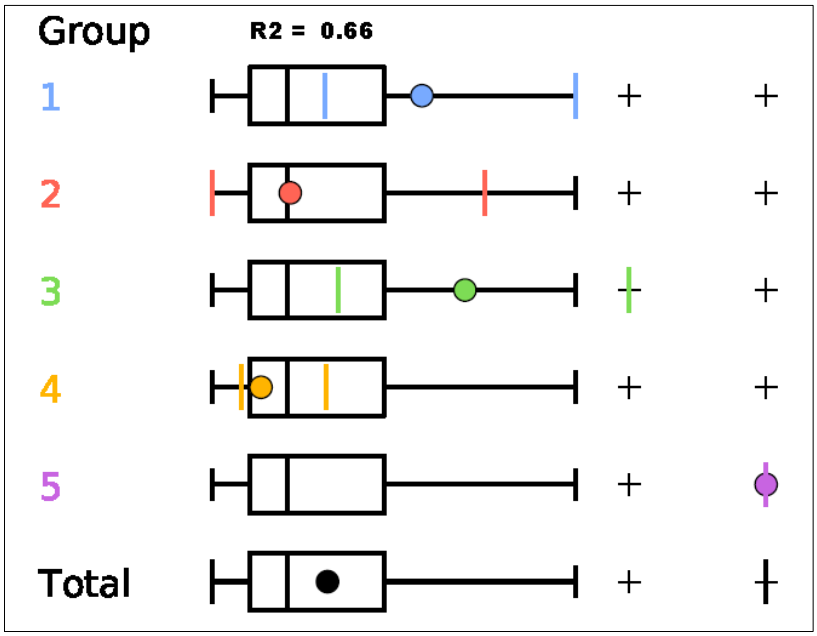

Fig. 14 Health Needs Statistics

The parallel box plot graph (Fig. 16) summarizes all the groups and the variables within them. From the graph, notice that group 2 (red) is close to median values other groups such as group 01, 03, 04, 05 are in the outlier part of the global statistics.

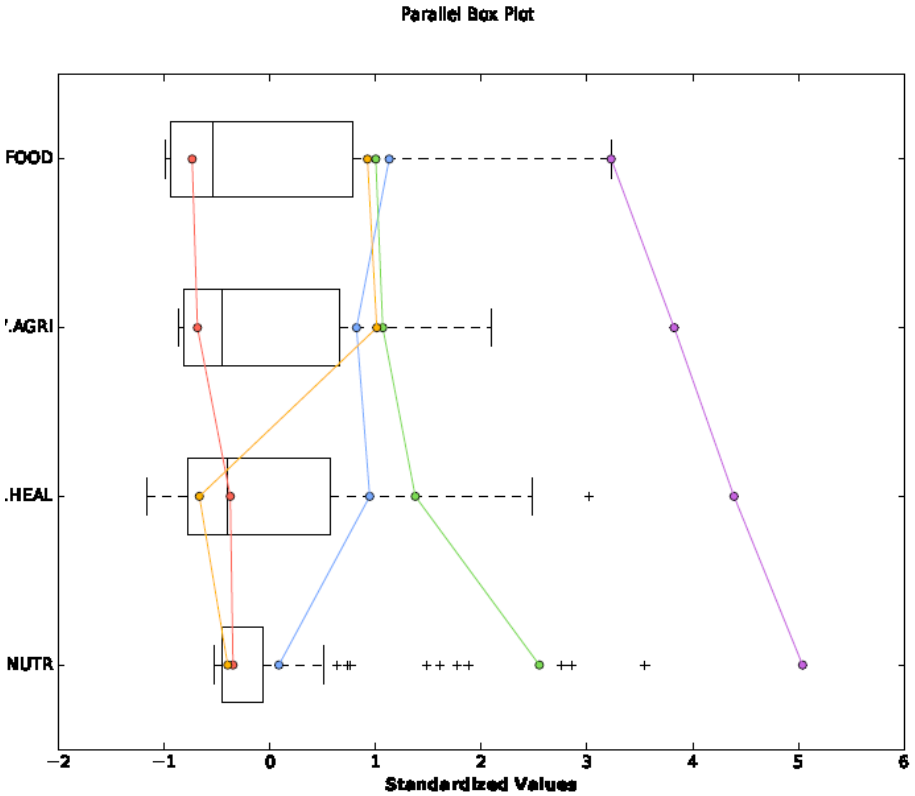

Fig. 15 Parallel Box Plot

4. Discussion

In this work five groups were identified and each group has its own specific feature. The major research question answered in the present work was to find the spatial and temporal variability in term of needs at zone level in Ethiopia. The research explored the potential use of grouping analysis tool provided by ArcGIS platform for geospatial analysis. The findings will be helpful for all stakeholders working to minimise the food problem.

The evidence from this study points towards the those zones where problem is more. These results suggest that large numbers of the zones in Ethiopia are having that much problem, so to control the problem it is very important to take the serious measures on food supply. So In this paper we have obtained satisfactory and comprehensive results proving international migration as an important factor the health issues. Our study provides the framework for a new way to use the GIS for the spatial/grouping study of need problems. These findings add to a growing body of literature on our understanding of GIS. The present findings might help to solve the regionalisation criteria issues based on certain factors. Our technique could be applied to other fields of spatial sciences also to group the places based on factors. We hope that our research will serve as a base for future studies on grouping analysis.

Acknowledgement

We thank the OCHA initiative to provide up to need data.

Conflict of interest

The authors declare no competing financial interests.

References

ArcGIS Desktop. (2018). How Grouping Analysis works., from Arcgis.com website: https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/how-grouping-analysis-works.htm

Bennett, L. (2018). Machine learning in ArcGIS. from https://www.e-education.psu.edu/geog858/sites/www.e-education.psu.edu.geog858/files/Lesson_04/Files/Machine-Learning-in-ArcGIS.pdf

Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically Weighted Regression. Journal of the Royal Statistical Society: Series D (the Statistician), 47(3), 431–443. https://doi.org/10.1111/1467-9884.00145

Ethiopia | Humanitarian Assistance | U.S. Agency for International Development , from Usaid.gov website: https://www.usaid.gov/humanitarian-assistance/ethiopia

Kröner, Alfred , Smedley, Audrey , Clarke, John Innes , Steel, Robert Walter , Mabogunje, Akinlawon Ladipo , Middleton, John F.M. , Gardiner, Robert K.A. , McMaster, David N. , Dickson, Kwamina Busumafi and Nicol, Davidson S.H.W.. "Africa". Encyclopedia Britannica,

Lyons, Terrence (1996). "Closing the Transition: The May 1995 Elections in Ethiopia". Journal of Modern African Studies. 34 (1): 135. doi:10.1017/S0022278X00055233

Murack, J. (2017). Regression Analysis Using ArcMap. Retrieved from website: https://libraries.mit.edu/files/gis/regression.pdf

OCHA. (2017). Ethiopia Severity of needs analysis - Ethiopia 2017 Severity of needs analysis.xlsx - Humanitarian Data Exchange, from Humdata.org website: https://data.humdata.org/dataset/ethiopia-2017-severity-of-needs-analysis-xlsx/resource/5837b491-f34a-40a7-8bb0-135330aa88da?view_id=d5bbcd0b-1867-4739-a93e-2b0fe886c6f4

USAID (2018) Fighting Malnutrition in Ethiopia Through Mobile Health and Nutrition Teams, from website: https://www.usaid.gov/sites/default/files/documents/1866/06.14.18

World Bank. (2017), from Worldbank.org website: https://datacatalog.worldbank.org/dataset/world-bank-official-boundaries/resource/12cc761c-bb67-45a7-82ff-eb64b67bdbdc