

Visualizing Need: Using Data Visualization as a Decision Support Tool in the Formulation of Social Policy

Randy Basham

University of Texas at Arlington, Texas, United States

International social work is facing an exponentially growing number of critical challenges in developing nations in meeting essential social needs as well as the distribution of material goods, social and economic resources and the provision, and effective services delivery to populations of need. There is a need to respond rapidly to populations in need and adapt to quickly changing socio-economic conditions. However, the resources needed to evaluate data and make effective decisions to provide support are often limited by the availability of trained data analysts who can interpret statistical findings and make relevant service delivery decisions for such communities of need. Modern computing technology shows promise of providing a graphical evaluation approach to enriching understanding of the available information. Thereby, the need to make critical decisions is augmented for those decision-makers that are working most closely with marginalized groups. A recent research effort undertaken by the Adventist Development and Relief Agency (ADRA) sought to provide administrators and policy-makers with information regarding the associations between social capital, economic development, and food security in the mountain region of Perú. Survey data collection and descriptive statistical analysis were used to provide information on the relationship between socio-economic development services and the emergence of social capital needed to develop effective human self-support systems. This paper demonstrates how to provide a graphical interface between areas of economic need and development initiatives. A visual exploration of data patterns would serve to inform human service decision-makers as they develop policies and programs. Visualization of data helps to enrich meaning and understanding of core areas of population need. This enhanced information format can identify needs for social policy development and improve the delivery of critical services and resources. This paper will provide a conceptual argument for considering data visualization methods in the evaluation of significant needs for, at risk, international populations. The paper will also present a number of examples for improving data evaluation and the understanding of need through data visualization methods through widely available computing technology. Procedures for graphically representing data will be demonstrated using the Perú survey, and applications for social policy formulation will be discussed.

Keywords: data visualization techniques, inform social policy, program development

Introduction

The developing world and Latin America in particular, are experiencing great economic, nutritional, and social support needs. Because of this, hundreds of non-governmental organizations collaborate with national, regional, and local governments in an effort to promote socio-economic development. These efforts, however,

are often hampered by the lack of empirical findings and decision-support information that can justify interventions. The lack of such empirical supports is often exacerbated by a shortage of well-trained researchers and development professionals who can bridge the gap between theory, research, and practice. Sophisticated data analysis and interpretation skills are usually needed to inform decision-making in development work as well as the social policies that should guide and justify such action. The previously-described situation challenges social work researchers to identify new and more effective ways to use research and technology to inform decision-making.

New techniques have recently emerged that use widely available computer technology to graphically illustrate data in highly sophisticated and informative ways. These data visualization techniques, some of which will be illustrated in this paper, involve a spectrum of graphical representation procedures that allow us to explore complex data sets and derive greater meaning from them. Numerical information reported in tabular form may not provide the optimal medium for exploring data trends and conceptualizing policy needs often aggregated in large data sets.

To demonstrate the advantages of data visualization in evaluating need, the author has performed a secondary analysis of a survey data set. This effort dramatically illustrates the ease of conveyance of meaning derived from numerical data using graphical representation as a data exploration technique. The initially analyzed survey data was collected in 1999 as a portion of social and economic needs study conducted in Perú, South America. The research aimed at exploring the associations between social capital, economic development, and food security in a highly impoverished region of the country (Dáz, Drumm, Ramírez-Johnson, & Oidjarv, 2002). The research evaluated the efforts of the Adventist Development and Relief Agency (ADRA) in Perú to meet social, economic, and nutritional need in 10 experimental communities while comparing needs and efforts to 10 control communities.

Adventist Development and Relief Agency (ADRA)

The Adventist Development and Relief Agency (ADRA) in Perú operates the Andean Food Security Development Project with USAID Title II funds. The project, which initially started in 1996, comprises the Infant Nutrition and the Agricultural Income Generation programs. These programs have targeted areas of extreme poverty in 13 of the country of Perú's political departments.

The project's qualitative midterm evaluation was completed on 22 January 1998 reported that ADRA had performed outstandingly in the area of community confidence building and mobilization (ADRA, 1998). Furthermore, the ADRA/Perú leadership strongly believed that repairing social relations and mobilizing communities in a country torn by years of violence and natural disasters is a critical preliminary step for further development. There was subsequently a need to critically evaluate relief and support effort effectiveness.

For these reasons, in 1999, ADRA/Perú commissioned the original study with the purposes of: (1) quantitatively testing the findings of the mid-term evaluation; (2) testing the validity of social capital theory for non-western cultures; and (3) evaluating the impact of socio-economic development services on an experimental ($N = 10$) and a comparison group ($N = 10$) of communities. The surveyed communities are located in Perú's mountain region. ADRA operates most of its programs in the mountain region because of its well-documented high incidence of poverty (Pillai, Diaz, Basham, & Ramírez-Johnson, 2011).

High Poverty in Perú's Mountain Region

A history of oppression, racial discrimination and inequality, legal barriers, adverse weather conditions, and insufficient means of communication and transportation are factors that have contributed to the high incidence of poverty in Perú's mountain region. Therefore, there is a substantial social need and a subsequent need to evaluate economic and policy efforts aimed at providing relief. Methods to evaluate need that serves to provide decision support for policy initiatives are critical.

Socio-economic inequality in Perú dates back to the Inca Empire before the arrival of the Europeans. Forced labor and extremely high taxes in the Inca Empire led to poverty and social stratification (Gisbert, 1999). This situation worsened with the arrival of the Spaniards who imposed a system of institutionalized discrimination against Mestizos whom they perceived as morally impure (Cadena, 1997). The arrival of African slaves who were mostly settled in the mountain region, and the subsequent arrival of Asians contributed to the complexity and inequality of the Peruvian social stratification system (Gisbert, 1999).

Crabtree and Thomas (1999) identified other factors contributing to the region's high level of poverty. First, prior to 1923 residents of the mountain region were not considered legally Peruvian, could not own property and lacked legal and social protections. Secondly, this area is mostly populated by peasants who are at the bottom of the socio-economic strata.

Adverse weather, such as drastic temperature changes, little rain, and severe erosion makes farming extremely difficult, and in this way contributes to poverty (Gisbert, 1999). Besides, from 1979 to 1992, the Peruvian government did not provide needed protections to the country's agriculture (Ferrary, 1992). According to Espinoza-Uriarte, (1997, p. 36) and Crabtree and Thomas (1999), life in the mountain region is made more difficult by inadequate housing, lack of clean drinking water, and insufficient means of transportation and communication (Crabtree & Thomas, 1999).

The Need for Decision Support in Social Policy Formulation

Social scientists and development professionals, such as those exploring need in Perú, need to know how social capital relates to economic development and food security since such a finding would have direct and significant implications for the theory and practice of development. Faced with the reality of limited financial resources, professionals often need to make difficult and strategic decisions related to what programs to implement first and where. They often struggle with questions, such as:

1. Should development programs be started only in communities that possess high levels of social capital in an effort to enhance the likelihood of success and sustainability?
2. Should economic development programs be implemented first in order to increase social capital?
3. Should social capital be increased through alternative means prior to implementing economic development programs?
4. Should social capital be viewed as a precondition for development or an outcome of economic development projects that must be monitored and assessed?

Policy-makers and development program administrators often struggle to agree on how, when, and where to allocate scarce resources. Decisions aimed at developing social capital, economic development, and community infrastructure are dependent on the collection and organization of data and information. New computer programs and applications are playing an increasingly important role in data collection and analysis

needed to inform decision-making. Technological supports are especially crucial when the data available to inform decisions is highly complex given that this makes the deductive and inferential process more elusive.

Recent trends in social service practice emphasized the need for evidence-based decision-making in human service delivery. This type of decision-making is especially important during times and in locations of economic uncertainty. The poverty and human need of the mountainous regions of Perú can rapidly take advantage of evaluation methods that explore data trends and provide understanding.

Data Visualization in the Determination of Need

Data visualization refers to a range of graphical representation procedures used with complex data sets that are often augmented by computer technology (Yu, 1995). The application of data visualization techniques show promise for increasing understanding and effectively enhancing decision-making. Computer support is necessary for creating graphics from large data sets, while historically handmade images or graphics may have sufficed to determine need when working with smaller data sets (J. W. Tukey & P. A. Tukey, 1988; Tukey, 1988a).

Data visualization methods make it possible to rapidly evaluate significant needs and identify at risk, vulnerable, or marginalized populations. These methods provide visual feedback and improved understanding of data available to guide policy formulation. Furthermore, the visual feedback is reinforced by available data and made cost-effective through the use of commonly available computing technology.

Image Development Methods to Assess Need

Commonly available computer spreadsheet applications have been demonstrated to have the capacity to collect and analyze complex data sets and generate graphical representations of complex data (Haber, 1988). Recent advances indicate that data visualization methods may be employed to capture process and outcome variables simultaneously and to compare change across multiple units of analysis and in multiple groups over time (Patterson & Basham, 2002; Basham, 2002). Furthermore, data visualization methods have the capacity to be used to convey meaning and enhance understanding across domains or levels of human service organizations (Patterson & Basham, 2003). Graphical representations may be matched to specific research questions so as to provide meaningful information for specific interests, including multivariate research questions and those that would address multiple areas of need (Yu, 1995) and as decision support for evidence-based practice in governmental service delivery systems (Schoech, Basham, & Fluke, 2006).

The creation of simple to interpret graphics reported along with the critically associated statistical information can improve data evaluation in the understanding of social and economic need (Tukey, 1988b). Graphical representation of relevant data sets may be accomplished through widely available computing technology through the vehicle of spreadsheet programs. Data may be analyzed by the use of traditional data analysis functions of a spreadsheet program, such as those incorporated into Microsoft Excel programs and through use of a robust data analytic pivot table technology incorporated into the program. The resulting numerical information can then be graphed using the chart wizard function of the program to produce readily interpretable graphical representations of the data. The procedures for graphically representing data include entry of data into a spreadsheet application and determination of the variables that may be useful for representation through a particular graphical display. Thereafter, variable data are selected and entered into the pivot table function of the spreadsheet application. Tabular information entered into the pivot table may then be exchanged as needed with other variables of interest. To create any number of pivot table configurations of

relevant data. The pivot table function quickly allows the evaluator to recalculate data so that sums, counts, averages, standard deviations, variation, and minimum and maximum data range may be selected. Once selected, the data may be formatted to graphically display by using functions of the chart wizard. Various graphs may be selected and sampled until a graph type with the best display properties for the data is finally chosen. Table 1 demonstrated the tabular data derived from the pivot table application to categorize high and low levels of social capital within 20 communities surveyed in Perú.

Communities with high social capital totaled 44.43 across 11 communities and communities with low social capital totaled 26.14 across nine communities.

Table 1

Excel Pivot Table for High and Low Social Capital in 20 Communities

Frequency of food need (past 12 months)		Low or high social capital	
Community	1.00	2.00	Grand total
1.00	3.27		3.27
2.00		2.22	2.22
3.00	4.53		4.53
4.00		1.98	1.98
5.00		3.13	3.13
6.00	3.46		3.46
7.00	3.05		3.05
8.00	4.29		4.29
9.00		3.52	3.52
10.00		1.45	1.45
11.00		5.96	5.96
12.00		2.38	2.38
13.00	4.65		4.65
14.00	2.74		2.74
15.00		3.53	3.53
16.00	3.02		3.02
17.00	3.74		3.74
18.00	6.51		6.51
19.00	5.18		5.18
20.00		1.96	1.96
Grand total	44.43	26.14	

The 3-dimensional bar graph depicted in Figure 1 readily illustrated that the 20 communities were nearly evenly split in reporting whether their community was perceived as having high (dark gray bars) or low (light gray bars) social capital. Columns on the bar graph numbered 1 through 10 represent the experimental communities receiving ADRA services. Columns 11 through 20 represent the comparison communities that did not receive ADRA services. The average number of times that respondents reported needing food, during the previous 12 months, is represented by the vertical axis. Community 11 reported both high social capital and a high incidence of respondents being able to obtain adequate food over the previous year. The incidence rate of hunger and the inability to secure adequate food was more, nearly twice as likely, in communities with low social capital as in the pivot table in Table 1.

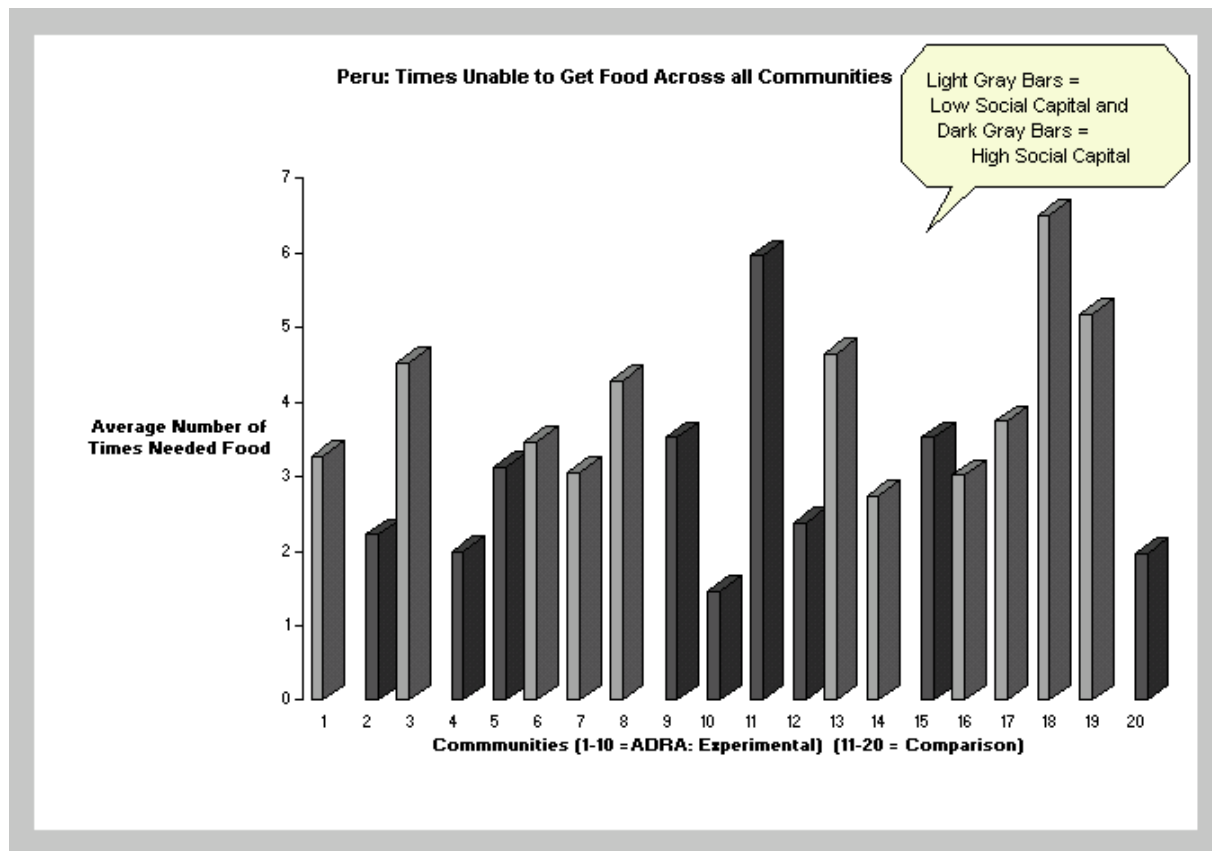


Figure 1. A 3-dimensional bar graph representing lack of food access in 20 communities.

In the pivot table depicted in Table 2, there are nine comparison communities where ADRA services were not provided, but where respondents reported residing. In these nine comparison communities, only three communities were identified as having low social capital (Communities 1, 4, and 9). Comparison communities with high social capital totaled 22.81 across communities and control communities with low social capital totaled 7.87 across communities.

Table 2

Excel Pivot Table for High and Low Social Capital in Comparison Communities

Comparison communities	1.00	2.00	Grand total
1	0.00	2.38	2.38
2	4.65		4.65
3	2.74		2.74
4	0.00	3.53	3.53
5	0.00		0.00
6	3.74		3.74
7	6.51		6.51
8	5.18		5.18
9	0.00	1.96	1.96
Grand total	22.81	7.87	30.69

The 3-dimensional bar graph depicted in Figure 2 readily illustrated that the comparison communities were somewhat split in respondents reporting whether their community was perceived as having high (dark

gray bars) or low (light gray bars) social capital. As in the pivot table in Table 2, one community did not report levels of social capital and associated instances of inability to obtain adequate food (Community 5), and so there is no corresponding bar for this community. In the comparison communities, low instance of hunger or inability to obtain food in the past year is reported for communities identified as having high social capital. Communities reporting low social capital had a higher number of reported instances of inability to obtain adequate food in the previous year.

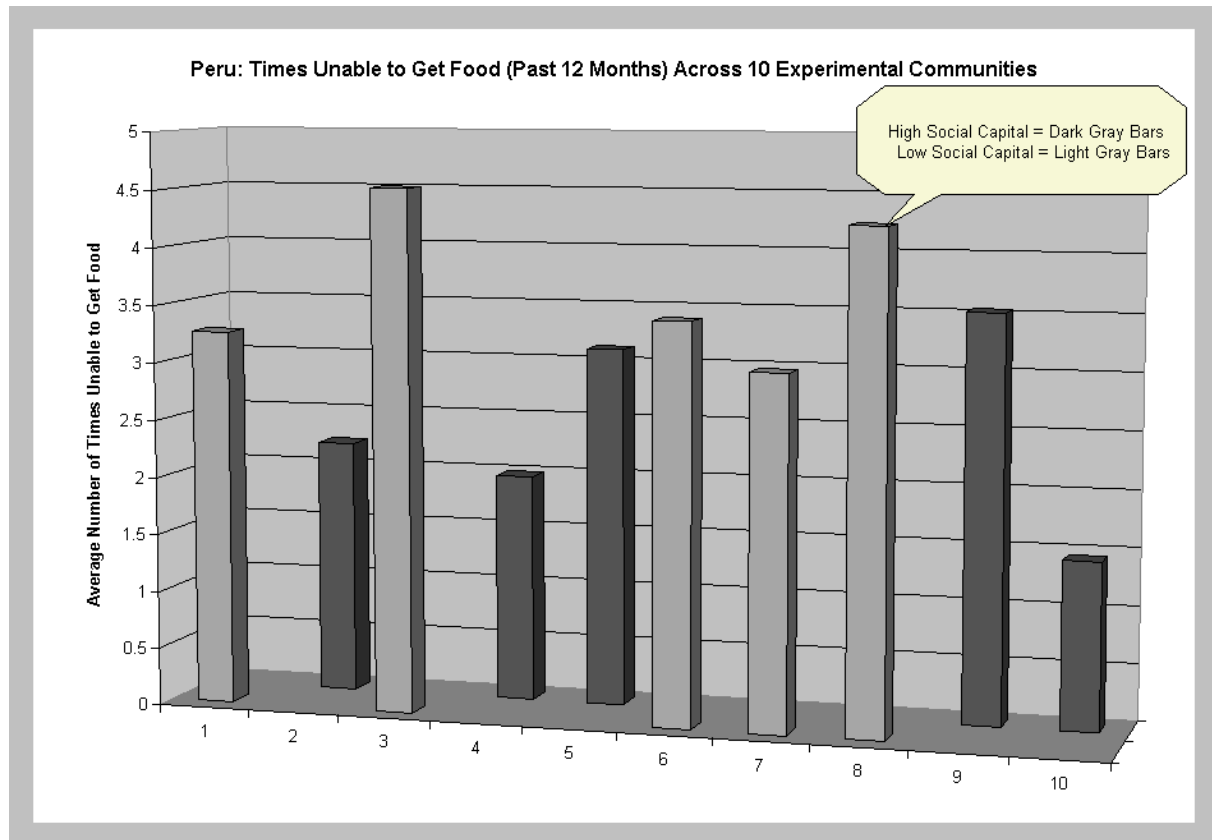


Figure 2. A 3-dimensional bar graph representing the lack of access to food in experimental communities.

Graphical Representations as Decision Support

Recent literature has demonstrated the capacity of line graphs enhanced by upper bound and lower bound standard deviations to detect agreement and turbulence in small groups (Patterson & Basham 2003). These have the potential of providing decision support for the progress and determining the conflict level within groups or as a supervisory tool for self-directed or supervisor directed work teams. Additional literature points to the guidelines for effective graphical use in the evaluation of group related services or interventions (Basham, 2002). The 3-dimensional surface plots through the addition of a third axis for displaying information derived from data can demonstrate information about more than one unit of analysis, such as individuals within groups and at the comparison of group level dimensions (Patterson & Basham 2002). Furthermore, data visualization has been considered as a method of providing decision support at multiple levels of human service agency need and more recently as an information feedback system for key decision makers within state government providing protective services to children (Schoech, Basham, & Fluke, 2006). However, the graphical displays

that would provide critical decision support from one administration of survey data, or for static population information that is not of a time series nature, as available on a particular population, and over the lifespan have been difficult to represent.

The pivot table in Table 3 provides aggregated information of a static or single survey nature for three continuous variables. These include age of the residents of the experimental communities, the education level of these residents, and the income level of the survey respondent.

Table 3

Excel Pivot Table for Age Education and Income in 10 Experimental Communities

Average of income	Grade								
Age	1-2	3-4	5-6	7-8	9-10	11-12	13-14	15-16	Grand total
18-22	1303.75	1627.69	1594.28	2550.00	1080.00	3280.00			1927.41
23-27	2437.50	1898.35	2012.52	2213.00	1771.69	1671.17	2400.00	2600.00	1994.72
28-32	1604.09	1698.71	1633.61	2551.87	2467.69	1664.00	250.00	652.00	1830.79
33-37	1538.46	1722.16	1788.33	2530.66	2747.50	2193.00		5390.00	2137.01
38-42	2091.08	1326.50	1731.37	3417.00	2993.33	1888.00	2400.00	11300.00	2182.06
43-47	754.77	2696.15	1394.00	1653.12	2671.25	1686.42			1769.19
48-52	1278.68	1211.68	1454.20	2842.22	350.00		2400.00	5600.00	1698.18
53-57	2512.00	1340.00	2388.00	1364.00		11500.00	3000.00		2284.40
58-62	1635.83	2038.00	1030.00	800.00	2100.00		7200.00		1633.75
63-67	1882.00	1225.00	630.00	2100.00	870.00				1339.41
68-72	5015.71	2700.00	1245.00	1440.00	2700.00		600.00		3139.37
73-77	2160.00		1920.00						2000.00
78-82	2617.50								2617.50
Grand total	1902.50	1695.63	1662.40	2507.08	2283.80	2152.36	2581.25	4891.50	1967.18

These three variables of age, education, and income level of survey respondent can be displayed to provide decision support for the possible allocation of resources, such as economic aid, social capital, and nutritional supports through the plotting of these three variables on a rarely used radar graph available in the spreadsheet application of Microsoft Excel programs. In Figure 3, age is plotted as the circumference of the outer circle from age of 18 to age of 81 years and constitutes a first axis of plotted continuous information. Then, concentric nested circles that radiate outward from the midpoint of the radar graph represent differing levels, or strata of potential income for survey respondents represented in soles, the currency of Perú. The range depicted for the 20 experimental counties is from zero to about 35,000 soles and represents the second axis of continuous data for survey respondents. Finally, the third axis of continuous data that are plotted as line graphs conforming to a central midpoint of the chart as the baseline point for no education, is the educational levels of respondents, depicted as lines, with the point values of the lines indicating the income level of the respondent and plotted at the radar location that conforms to the respondents reported age. By viewing the graph and following the circumference axis of age income levels are seen to gradually increase from age 18 to about age of 42 years and then decrease gradually until about age of 53 years. Thereafter, much of the elderly respondents are represented to have little education and are impoverished with the exception of a small group of respondents who are seen as outliers and are financially well of and well educated and are about 68 years of age. This graphic provides a great deal of unique information that captures the relationship between age and education to income level across an adult generation of respondents within the 20 experimental counties. This information provides valuable decision support in term of allocating nutritional programs and intensifying

educational opportunities. The graphic also painfully points out the plight of the devastating poverty affecting most respondents average age is 53 years old. The visual display of this complex information both explores the data and provides a rapid understanding of the interaction of age, education, and income.

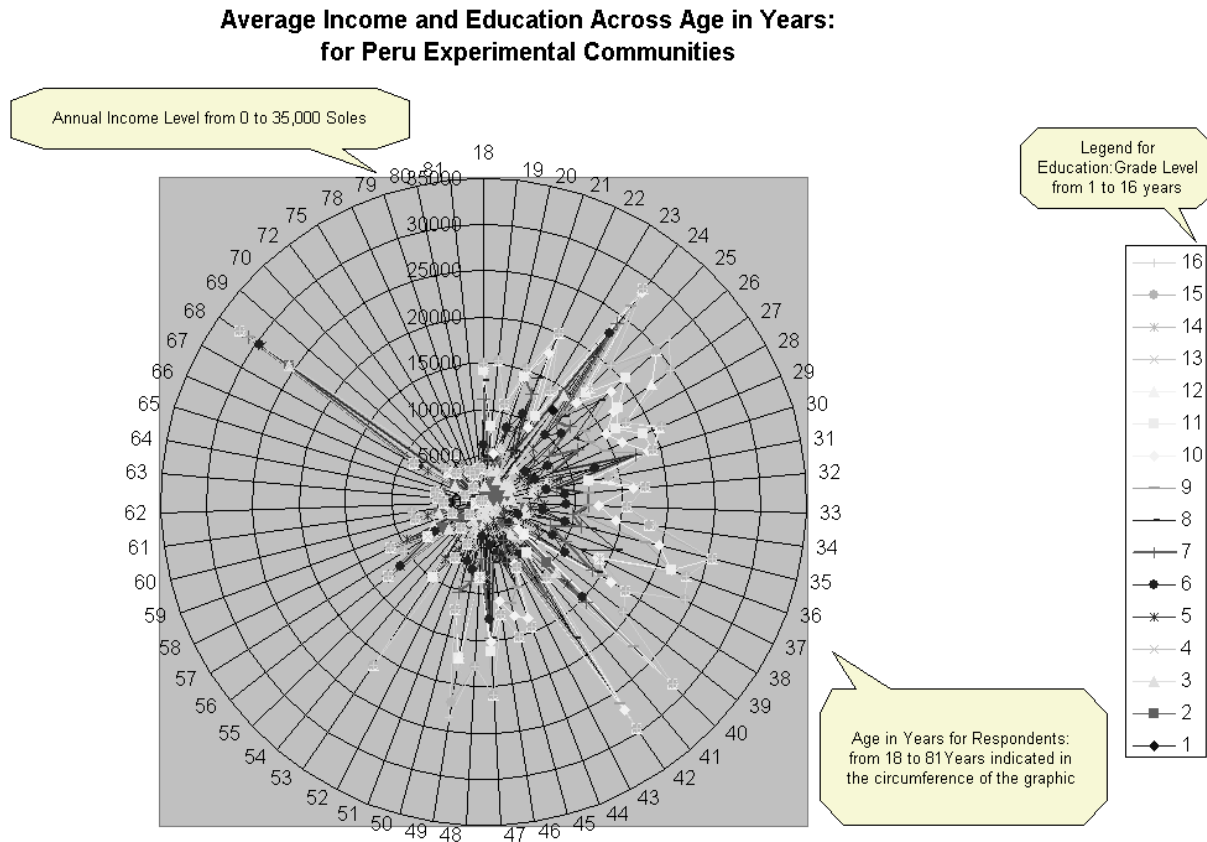


Figure 3. A “radar” chart for age education and income in 10 experimental communities.

Utility: The Visual Analysis of Data

Visual analysis of data only refers to the visual array of one set of information relative to another collection of information (Tufte, 1983). Graphical representations or arrays are preferred when at least one set of the information is quantitative. Pictures or diagram arrays are preferred when all of the sets of information are qualitative (Tufte, 1990). Many conventions have been proposed for interpreting data using a graphical representation that emphasize the provision of perceptual clues, such as grid marks, surface lighting, and varying texture, color, and orientation (Haber, 1988).

The utility questions of which data visualization approaches are the most useful with differing types of social policy data have not yet been well established. However, the usefulness of graphical representation in the ability to illustrate trends in complex data and provide meaningful decision support information for those who determine or evaluate the effectiveness of social policies is rapidly becoming an established convention.

Discussion: Data Visualization and Social Policy

The application of data visualization with respect to the utilization of graphical representation to inform

and guide decision-makers in the formulation of social policy is a relatively new concept (Cobb, 2003) and has been used primarily with geographic information systems to guide educational policy. In this study and case example, disparities in communities in terms of social supports and available resources may be rapidly evaluated using the visual representation of data. Decision- and policy- makers can interpret and understand these difference and design useful policy measures to reduce excessive need and poverty and to promote sharing behaviors and community self-help.

Conclusion

Readily available and robust data analytic tools exist in desktop software like Microsoft Excel that can provide decision support for community projects involving micro-economic development. Input data may be quickly analyzed using a range of tools that improve comprehension of complex social and economic issues through graphical representation. Once understood, these outputs can provide the basis for developing comprehensive social and economic policy. Equitable distribution of goods and services can be improved to those in greatest need in communities suffering from economic need or various times of food, shelter or other insecurities.

Community, group and family, or individual needs may also be analyzed and deficits addressed. The same data imaging and visualization processes could also be used in temporary service, economic, or goods disruptions in distribution, such as in small local negative or traumatic events and even in case of disaster's so long as adequate data in the population is either available or can be quickly and efficiently collected.

References

- Adventist Development and Relief Agency (ADRA). (1998). *Midterm evaluation: Andean food security development project*. Lima, Perú: ADRA.
- Basham, R. (2002). Data visualization: Graphical representation in the evaluation of experiential group therapy education outcomes (Doctoral dissertation, The University of Tennessee).
- Cadena, M. (1997). *La decencia y el respeto*. Lima Perú: I.E.P. ditions.
- Cobb, C. D. (2003). Geographic methods and policy: Using geographic information systems to inform education policy. *Educational Research Quarterly*, 27(1), 28-40.
- Crabtree, J., & Thomas, J. (1999). *The Peru of Fujimori*. Institute of Peruvian Studies.
- Díaz, H. L., Drumm, R., Ramírez-Johnson, J., & Oidjarv, H. (2002). Social capital, economic development and food security in Perú's Mountain Region. *The International Social Work Journal*, 45(4), 481-495.
- Espinoza-Uriarte, H. (1997). *Regional development: A preliminary approach*. Lima, Peru: National University Federico Villareal, University Press.
- Ferrary, C. A. (1992). *Industrialization and development: Public policies and economic effects in Peru*. Lima, Peru: Fundación Friedrich Ebert.
- Gispert, C. (1999). *Concise Encyclopedia of Peru*. Barcelona, Spain: Oceano Editorial Group.
- Haber, R. B. (1988). Visualization in engineering mechanics: Techniques, systems, and issues. *Proceedings of ACM Computer Graphics (SIGGRAPH'88)*, August 1-5, Atlanta, Georgia.
- Patterson, D. A., & Basham, R. E. (2002). Data visualization procedures in the analysis of group treatment outcomes across units of analysis. *Small Group Research*, 33(2), 209-33.
- Patterson, D. A., & Basham, R. E. (2003). Visualizing change: Spreadsheets and graphical representation across domains in human service practice. *Journal of Technology in Human Services*, 21(4), 1-16.
- Pillai, V., Diaz, D. H., Basham, R., & Ramírez-Johnson, J. (2011). Democratic attitudes and social capital in Latin America. *International Social Work*, 54, 767-779.
- Schoech, R., Basham R. & Fluke, J. (2006). Applying EBP (evidence based practice) in CPS (child protective services): A technology enhanced model. *Journal of Evidence-Based Social Work*, 3(3/4), 55-72.

- Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.
- Tufte, E. R. (1990). *Envisioning information*. Cheshire, CT: Graphics Press.
- Tukey, J. W. (1988a). *The collected works of John W. Tukey*. In W. S. Cleveland (Ed.), *Volume V Graphics 1965-1985*. Pacific Grove, CA: Wadsworth and Brooks/Cole Advanced Books and Software.
- Tukey, J. W. (1988b). What statisticians have been forgetting. In W. S. Cleveland (Ed.), *The collected works of John W. Tukey: Volume IV. Philosophy and principles of Data Analysis 1965-1986* (p. 5). Pacific Grove, CA: Wadsworth and Brooks/Cole Advanced Books and Software.
- Tukey, J. W., & Tukey, P. A. (1988). Computer graphics and exploratory data analysis: An introduction. In W. S. Cleveland (Ed.), *The collected works of John W. Tukey: Volume V. Graphics: 1965-1985*. Belmont CA: Brooks/Cole.
- Yu, C. H. (1995). The interaction of research goal, data type, and graphical format in multivariate visualization (Doctoral dissertation, Arizona State University).