

## Probing Needs Assessment Data in Depth to Target Programs More Effectively

### Abstract

Extension professionals often assess community needs to determine programs and target audiences. Data can be collected through surveys, focus group and individual interviews, meta-analysis, systematic observation, and other methods. Knowledge gaps are identified, and programs are designed to resolve the deficiencies. However, do Extension professionals look deeply enough into the data to identify subgroups that will allow more targeted programs, or are programs based on only superficial analysis? Cluster analysis allows data to reveal demographic patterns and relationships involving survey respondents. It allows Extension professionals to more precisely target program audiences and thus effectively achieve program impacts.

**JoAnne Skelly**  
Associate Professor  
and Extension  
Educator  
University of Nevada  
Cooperative Extension  
Carson City, Nevada  
[skellyj@unce.unr.edu](mailto:skellyj@unce.unr.edu)

**George Hill**  
Associate Professor  
Department of  
Educational  
Leadership  
Reno, Nevada  
[gchill@unr.edu](mailto:gchill@unr.edu)

**Loretta Singletary**  
Professor and Area  
Director  
University of Nevada  
Cooperative Extension  
Reno, Nevada  
[singletaryl@unce.unr.edu](mailto:singletaryl@unce.unr.edu)

University of Nevada,  
Reno

## Introduction

Extension professionals regularly assess community needs to identify and prioritize issues to address through programs. Primary data are collected through postal surveys, focus group interviews, individual key informant interviews, meta-analysis, systematic observation, and other methods (Bazik & Feltes, 1999; Malmshemer & Germain, 2002; Yang, Fetsch, McBride, & Benavente, 2009). Usually knowledge gaps are identified, and programs are designed to address these gaps to ultimately effect positive behavioral change. To direct programs more effectively, Extension professionals can use cluster analysis. Cluster analysis allows for a deeper examination of primary data to reveal demographic patterns and relevant relationships among respondents. It allows Extension professionals to specify and more finely target audiences for program delivery, a necessary factor in the face of reduced Extension resources.

"The first lesson of social marketing is that there is no such thing as targeting the general public" (Weinrich, 1999, p. 5). Extension professionals in an era of declining budgets must ensure that they target their programs as efficiently as possible in order to be both cost-effective and results-effective. They cannot deliver programs effectively with a scattershot approach. They must be precise

in selecting their audience. They should not develop programs based on what they think an audience needs to know. Instead, they should identify trends for specific audiences based on a deeper look at data collected in needs assessments. Previous research published in *Journal of Extension* has examined the use of cluster analysis to examine demographic factors that can affect different clientele subgroups, classify them into homogeneous groups, and identify trends within these groups (Mangiafico, Obropta, & Rossi-Griffin, 2012; Morford, Kozak, Suvedi, & Innes, 2006).

Clustering is a method of modeling that separates data into smaller groups or clusters. Essentially, members of a cluster have characteristics more similar to each other's than to those of members of other clusters. Cluster analysis is a tool that can be used to summarize hundreds of thousands of observations on several variables by finding groups within the data (Wallin, 2010).

Cluster analysis is an effective tool for data mining and can be used to determine if a group of study respondents can be clustered by traits. Everitt, Landau, and Leese (2009) describe cluster analysis as a structured way to identify clusters of groups sharing the same characteristics. It is used in a wide range of situations requiring exploratory data mining. It is a common technique for statistical data analysis used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, bioinformatics, and market research (Everitt et al., 2009).

## Methods

In 2008, the University of Nevada Cooperative Extension, in partnership with the University of Nevada, Reno, Latino Research Center (LRC), developed and implemented a statewide survey to assess Extension educational program needs of Latinos in Nevada. The goal of the assessment was to discover the primary concerns of Latinos living in Nevada that might be addressed by Extension programs. The researchers also wanted a generalized, but realistic picture of what life is like for Latinos in Nevada, to subsequently prioritize and tailor future outreach programs to meet these identified needs. Data collection and basic analysis was completed in 2009 (Skelly, Singletary, Angle, & Sepúlveda-Pulvirenti, 2010).

For the purpose of the study, the researchers designed a questionnaire to assess Latinos' perceptions of needs that might be addressed through Extension programs. The researchers piloted the questionnaire with stakeholders and volunteers from Nevada's Latino community and revised it based on their suggestions.

The resulting three-page questionnaire featured 32 issue items. Using a Likert-type scale of 1 (not very important) to 5 (very important), survey participants were asked to rate the importance of each issue. The questionnaire included a number of demographic questions such as age, gender, marital status, education, years lived in Nevada, occupation, occupation in country of origin, country of origin, immigration status, and language spoken at home. It included open-ended questions to determine familiarity with Extension programming, availability for participation in Extension programs, and preferences for receiving educational information.

Practice demonstrates that use of the focus group provides Latino populations with a comfortable means to express their opinions (Farner, Rhoads, Cutz, & Farner, 2005; Hobbs, 2004; Malek, 2002). Subsequently, the authors hired a bilingual and bicultural Latina facilitator to survey volunteer

participants using the written questionnaire combined with group interviews organized as open public meetings. Participants received a copy of the questionnaire in Spanish or English and an accompanying cover letter. This method for collecting primary data was approved by the University of Nevada, Reno IRB review board. The facilitator collected 1,015 surveys from 14 of Nevada's 17 counties.

These 1,015 completed questionnaires served as the data source for the statewide Latino study. Because some participants did not answer all survey items, the number of responses varied by survey item. Cronbach's coefficient alpha was used to estimate internal consistency of the 32 Likert-type scale items. The Cronbach's alpha score was high ( $r = .92$ ). The high score for instrument reliability demonstrates a high degree of consistency among the 32 items (Carmines & Zeller, 1979).

Cluster analysis was conducted to further classify data in order to determine if response patterns differed between survey respondent subgroups. Subgroups were identified by analyzing question response patterns.

K-means, non-hierarchical, clustering was selected due to the size of the data set and the desire to retain a limited number of clusters. The investigators were also seeking a centroid method of identification of clusters versus a large number of clusters, which would be difficult to classify for educational program targeting. In this method, the number of clusters is specified (forced) and we used centroids/clusters of 3, 4, 5, and 6. The calculations were completed using SPSS, and, in the process, the method calculates centroids for a trial set of clusters (in this case: 3, 4, 5 or 6), then places each item in the cluster with the nearest centroid, recalculates the centroid, and then reallocates the items to the closest centroid. This process of iterations continues until there are no more changes in the cluster membership (Antonenko, Toy, & Niederhauser, 2012).

After reviewing data from all the various cluster solutions, the investigators chose the three-cluster solution (the four-cluster solution had one cluster with only two members) and proceeded to validate the solution based on disaggregated means for the three cluster solution. The differences between clusters were confirmed using this method (Antonenko et al., 2012).

## Results and Discussion

Primary data collected allowed the authors to examine not only basic demographic data such as gender, age, and marital status, but also to explore issues such as country of origin, occupation in country of origin, current occupation, and immigration status. Understanding the historical as well as contemporary aspects of respondents' experiences helped to develop a more comprehensive understanding of Latino needs in Nevada. While the primary data indicated directions for Extension programs, a deeper look at the data using cluster analysis allowed for a more refined understanding of the Latino population surveyed to facilitate targeting specific program subgroups and more precise program delivery methods.

After conducting the cluster analysis and validation, it was determined that the three resulting clusters or subgroups were indeed classifiable. Members of each cluster responded to the 32 items on the survey more similarly than did members of the other clusters. Of the 1,015 respondents, 799 were placed in one of the three clusters. In the cases where placement was not made into a cluster,

the respondent typically had characteristics of two of the existing clusters.

Cluster 1 was the smallest group (68), with the following demographic characteristics: little formal education, manual labor occupations, lower wages, and the highest percentage of males in any of the cluster groups. Cluster 2 was the next largest group (299) and had the following demographic characteristics: the highest educational levels of any of the clusters, populated predominantly by women, and members were typically U.S. citizens or legal residents of the country. Cluster 3 was the largest group (432). Respondents were primarily women, with a lower educational level than that found among the respondents grouped in Cluster 2, but higher than that found in Cluster 1. In addition, this group reported the most diverse resident status, in that many held special visa status, while almost no members of other clusters held special visa status. In all clusters, the country of origin was predominantly Mexico.

Subsequently, one-way analysis of variance (ANOVA) was then conducted using the three clusters as the grouping/demographic variable. ANOVA conducted using the three clusters as demographic variables showed a statistically significant difference in responses to all but two of the 32 questions on the survey. Post-hoc analysis showed that, in most cases, the differences were significant between each of the three clusters. Table 1 displays the metrics for four of the 32 items when ANOVA was conducted using the clusters as the grouping variable.

**Table 1.**  
Results of ANOVA on Selected Survey Items Using Clusters as  
the Grouping Variable

	df	ms <sup>2</sup>	F	Sig
<b>Speaking and Reading English</b>	<b>3</b>	<b>21.70</b>	<b>69.57</b>	<b>&lt;.001</b>
Error	796	.358		
<b>Using Computers</b>	<b>3</b>	<b>3233.0</b>	<b>268</b>	<b>&lt;.001</b>
Error	796	9700		
<b>Good Credit</b>	<b>3</b>	<b>99.06</b>	<b>193</b>	<b>&lt;.001</b>
Error	796	.458		
<b>Dropping out of School</b>	<b>3</b>	<b>259</b>	<b>8.884</b>	<b>&lt;.001</b>
Error	796	409		

The second set of metrics (Table 2) used educational level as the grouping variable in a one-way ANOVA. These items were selected for illustrative purposes only and are the same as shown in Table 1 for comparative purpose. Using the traditional demographic information as grouping variables resulted in a greatly reduced number of survey items where differences were detected.

**Table 2.**  
Results of ANOVA on Selected Survey Items Using a

Demographic Variable (Educational Level) as the Grouping  
Variable

	df	ms <sup>2</sup>	F	Sig
<b>Speaking and Reading English</b>	<b>8</b>	<b>.620</b>	<b>1.414</b>	<b>.188</b>
Error	910	.439		
<b>Using Computers</b>	<b>8</b>	<b>1.154</b>	<b>1.417</b>	<b>.185</b>
Error	894	.814		
<b>Good Credit</b>	<b>8</b>	<b>.776</b>	<b>1.218</b>	<b>.285</b>
Error	895	.637		
<b>Dropping out of School</b>	<b>8</b>	<b>2.579</b>	<b>1.319</b>	<b>.230</b>
Error	857	1.955		

Table 2 shows the results for the same survey items when ANOVA with educational level as the grouping variable. Similar results were found when each traditional demographic variable was used as the grouping variable.

The difference between the results illustrated in the two tables is that Table 1 shows the results of an ANOVA using the clusters as the grouping variable and Table 2 shows the results of one demographic variable with the same four survey items. These results are indicative of our findings when ANOVA utilized the clusters as grouping variables and when testing survey items using the eight individual demographic items from the survey as grouping variables.

These data appear to indicate that real differences are likely missed when relying on traditional demographic grouping variables. This means that programming needs and opportunities will be missed as well. Cluster analysis is not a difficult analysis to conduct when using statistical software. Extension professionals are strongly urged to consider using it regularly when analyzing survey data involving large numbers of observations. As the data collected for this assessment illustrate, the differences are statistically significant. This analysis suggests that subsequent programs that are tailored to groups, based upon their educational attainment, gender, age, and occupation, for example, are more likely to have real impacts.

## Conclusions and Recommendations

Cluster analysis provided additional and important insight into the data collected in the Nevada Statewide Latino Needs Assessment. It allowed the authors to identify and examine a set of subgroups from the overall respondent group to explore the potential for more tailored and targeted program delivery. Results confirmed that programming based on traditional grouping variables may lead to the needs of some participants not being met.

Large groups surveyed are generally more diverse than superficial examination of basic frequencies

and typical disaggregation of data suggests. Demographic constructs typically exist within the data that can provide a much more detailed picture of target clientele. Cluster analysis allows for that detailed picture to become clearer and subsequent programs tailored to be more effective.

## References

- Antonenko, P. D., Toy, S., & Niederhauser, D. S. (2012). Using cluster analysis for data mining in education technology research. *Journal: Educational Technology Research Development* . 60(3):383-398. doi:10.1007/s11423-012-9235-8.
- Bazik, M., & Feltes, D. (1999). Defining your customer profile - an essential tool. *Journal of Extension* [On-line], 37(6). Available at: <http://www.joe.org/joe/1999december/a4.php>
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment* . Beverly Hills, CA: Sage Publications.
- Everitt, B. S., Landau, S., & Leese, M. (2009). *Cluster analysis* (4th ed.). London, England: Arnold Publishers.
- Farner, S., Rhoads, M. E., Cutz, G., & Farner, B. (2005). Assessing the educational needs and interests of the Hispanic population: The role of Extension. *Journal of Extension* [On-line], 43(4). Available at: <http://www.joe.org/joe/2005august/rb2.php>
- Hobbs, B. B. (2004). Latino outreach programs: Why they need to be different. *Journal of Extension* [On-line], 42(4). Available at: <http://www.joe.org/joe/2004august/comm1.php>
- Malek, F. (2002). Using the focus group process to assess the needs of a growing Latino population. *Journal of Extension* [On-line], 40(1). Available at: <http://www.joe.org/joe/2002february/tt2.php>
- Malmsheimer, R. W., & Germain, R. H. (2002). Needs assessment surveys: so they predict attendance at continuing education workshops? *Journal of Extension* [On-line], 40(4). Available at: <http://www.joe.org/joe/2002august/a4.php>
- Mangiafico, S. S., Obropta, C. C., & Rossi-Griffin, E. (2012). Demographic factors influence environmental values: A lawn-care survey of homeowners in New Jersey. *Journal of Extension* [On-line], 50(1). Available at: <http://www.joe.org/joe/2012february/rb6.php>
- Morford, S., Kozak, R., Suvedi, M., & Innes, J. (2006). Factors affecting program evaluation behaviours of natural resource Extension practitioners—Motivation and capacity building. *Journal of Extension* [On-line], 44(3). Available at: <http://www.joe.org/joe/2006june/a7.php>
- Skelly, J. A., Singletary, L., Angle, J., & Sepúlveda-Pulvirenti, E. (2010). *Addressing the needs of Nevada's growing Latino population: Results of a statewide needs assessment* . Special Publication 10-08. Retrieved from: <http://www.unce.unr.edu/publications/files/cy/2010/sp1008.pdf>
- Wallin, J. (2010). *Customer segmentation using cluster analysis on student loan applications* . (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1486963).
- Weinrich, N. (1999). *Hands-on social marketing: A step by step guide* . Thousand Oaks, CA: Sage

Publications Inc.

Yang, R. K., Fetsch, R. J., McBride, T. M., & Benavente, J. C. (2009). Assessing public opinion directly to keep current with changing community needs. *Journal of Extension* [On-line], 47(3). Available at: <http://www.joe.org/joe/2009june/a6.php>

---

*Copyright* © by Extension Journal, Inc. ISSN 1077-5315. Articles appearing in the Journal become the property of the Journal. Single copies of articles may be reproduced in electronic or print form for use in educational or training activities. Inclusion of articles in other publications, electronic sources, or systematic large-scale distribution may be done only with prior electronic or written permission of the *Journal Editorial Office*, [joe-ed@joe.org](mailto:joe-ed@joe.org).

If you have difficulties viewing or printing this page, please contact [JOE Technical Support](#)