Using Self Organizing Maps to Analyze Demographics and Swing State Voting in the 2008 U.S. Presidential Election

Paul T. Pearson[1] and Cameron I. Cooper[2]

[1] Hope College, PO Box 9000, Holland, MI 49422, USA
`paultpearson@gmail.com`
[2] Fort Lewis College, 1000 Rim Drive, Durango, CO 81301, USA
`Cooper_C@fortlewis.edu`
`http://faculty.fortlewis.edu/Cooper_C`

Abstract. Emergent self-organizing maps (ESOMs) and k-means clustering are used to cluster counties in each of the states of Florida, Pennsylvania, and Ohio by demographic data from the 2010 United States census. The counties in these clusters are then analyzed for how they voted in the 2008 U.S. Presidential election, and political strategies are discussed that target demographically similar geographical regions based on ESOM results. The ESOM and k-means clusterings are compared and found to be dissimilar by the variation of information distance function.

Keywords: Kohonen self organizing map, k-means clustering, variation of information, United States election 2008, United States Census data 2010.

1 Introduction

The United States presidential election in 2008 had many so-called swing states in which the election results were too close to predict accurately before election day and the margin for victory was narrow. Because of the close relationship between demographics and voting tendencies [2, 5–7, 13, 14, 20], this article examines the relationship between demographically similar counties and their voting tendencies for the 2008 U.S. presidential election in the three swing states with the most electoral votes (Florida, Pennsylvania, and Ohio). Emergent self-organizing maps (ESOMs) and k-means clustering were used to cluster demographic data provided by the 2010 U.S. Census, thereby identifying geographic regions (in this case, counties) that have similar demographics. After clustering, the voting results for the 2008 United States presidential election were examined within each each cluster of demographically similar counties. Sometimes, demographically similar counties in the same cluster voted for different candidates in the 2008 presidential election. The demographic clusters from ESOMs with mixed voting outcomes were examined closely, and it is suggested that a political party may be able to improve its chances for winning future elections

by applying strategies that worked for one county in a cluster with mixed voting outcomes to other counties in the same cluster. The results obtained using ESOMs and k-means clustering to cluster census data for each of these three swing states are compared, and variation of information distance calculations were used to measure the dissimilarity between ESOM and k-means results [15]. ESOMs and k-means clustering were chosen to highlight the differences between ESOMs, which have thousands of weights (or neurons), and k-means clustering, which is believed to yield results similar to ordinary SOMs which have only tens or hundreds of weights [23].

2   Background

Self organizing maps were created by Finnish professor Teuvo Kohonen in the 1960s. A comprehensive mathematical description of Kohonen's work can be found in his book *Self-Organizing Maps* [10]. A key feature of a SOM is that it produces a low-dimensional picture (usually two- or three-dimensional) of a high-dimensional data set in such a way that points near each other in the low-dimensional picture come from points that are near each other in the high-dimensional data set. This dimension reduction feature of a SOM has been very useful to researchers interested in visualizing clusters in high-dimensional data sets.

SOMs have been used for a wide variety of applications, such as bioinformatics, health care, finance, language processing, document analysis, and image processing [12]. In a paper closely related to this one, Niemelä and Honkela used SOMs to explore the relationship between four socio-economic factors (cost of living, unemployment, gross domestic product, and total consumption) for the entire country of Finland between 1954 and 2003 and parliamentary election results that involved nine political parties during that time period [18]. In contrast, this study uses ESOMs to cluster counties in three states of the United states based on 51 different socio-economic factors measured in the 2010 Census and how they are related to the 2008 presidential election results between essentially two political parties (Democrats and Republicans). In work similar in spirit to this, Kaski and Kohonen used SOMs to study the socio-economic status of the countries in the world based on World-Bank data [9]. Tuia, et al., have used a SOM for the clustering of urban municipalities in Switzerland depending on their socio-economic profile [22]. The mechanism for self-organization in a SOM has even been used in a non-computational way as a metaphor to explain the patterns in electoral processes [16, 17]. The second author has used SOMs to identify benchmark universities on the basis of student assessment of university websites [4].

3   Methods

This section explains the data acquisition and preprocessing, how self-organizing maps were used to cluster the data, and how k-means clustering was used to cluster the data.

3.1   Data Acquisition and Preprocessing

The states of Florida, Pennsylvania, and Ohio were chosen for analysis because they are states in which the 2008 U.S. presidential election was close [3], they have the highest number of votes in the electoral college among all swing states (see Table 1), and they are projected to be states won by very narrow margins in the 2012 U.S. presidential election [1].

Table 1. Electoral college votes for three states in the 2008 and 2012 U.S. presidential elections

	Electoral College Votes		
Year	Florida	Pennsylvania	Ohio
2008	27	21	20
2012	29	20	18

Demographic data for all of the counties in these three swing states were obtained from the U.S. Census Bureau website for the year 2010 census [24]. The census data includes population demographics by age, ethnicity or race, education level, housing data, income data, employment data, trade data, government spending data, and land area data. A complete list of all 51 census data variables can be obtained from the Census Bureau website [25]. One data set was created for each of the three swing states from the 2010 census data, and the observations were the counties in the state and the variables were the 51 census data variables. Since the variables were not uniform in scale, each variable was normalized using a z-transform to make it scale invariant.

3.2   Clustering by Emergent Self-Organizing Maps

Emergent self organizing maps (ESOMs) are SOMs with several thousand weights (i.e., neurons), whereas ordinary SOMs have tens or hundreds of weights. ESOMs have been shown to be significantly different from ordinary SOMs, and ordinary SOMs can yield results similar to k-means clustering [23]. Thus, ESOMs were chosen for comparison to k-means clustering to determine whether ESOMs yield different clusterings than k-means does. The authors created ESOMs using the Databionic ESOM Analyzer [23], which uses the standard SOM algorithm for weight updates [10, 11]. The ESOMs created had 4,100 weights that were distributed on 50×82 toroidal grids in order to avoid error effects that occur near the edges of rectangular maps [23].

Distances between observations and weights in ESOMs were calculated using Pearson distance $1 - \rho$, where ρ is the Pearson correlation between an observation and a weight. Correlation was chosen for distance measurements instead of Euclidean distance because Euclidean distance may give undue influence to one particular variable. For example, the 2010 population of Philadelphia county

is disproportional to the other counties in Pennsylvania, which is reflected by the fact that the z-score for Philadelphia county is 4.98 for the 2010 population variable. If Euclidean distance had been chosen for distance measurements, then very populous counties such as Philadelphia county would be separated by a sizable amount from counties with an average population, which would have a z-score near 0; hence, correlation distance was chosen to mitigate such effects. Also, using correlation tended to produce many clusters of small to medium size, whereas using Euclidean distance produced a few very large clusters and a smattering of very small clusters.

The topological ordering provided by the grid of an ESOM ensures that similar data points will be displayed in contiguous locations. However, by insisting on contiguous locations, the ESOM display suppresses the variation in the degrees of dissimilarity among the weights [21]. For this reason, a unified distance matrix, or U-matrix, was used to show inter-cluster distance in the ESOMs. The U-matrix was used to generate a contour map that visually separated clusters in the ESOMs.

The ESOMs were constructed using the following parameters. The number of training epochs was 30. The learning rate started at 0.95 and decreased linearly to 0.01. A two-dimensional output map grid on a torus was used with Euclidean distance on the grid. The other parameters used in the Databionic ESOM were the default values: online training was used, k for k-batch was 0.15, the initial map size was 10%, the ending epoch for good initialization was the 15th epoch, best matches were found using the standard search and a radius of 8, weights were initialized by a Gaussian distribution, the radius started at 24 and decreased linearly to 1, the neighborhood kernel function was a Gaussian, and the data patterns were permuted.

The ESOMs were configured to label each point with its county name and also color it according to how it voted in the 2008 U.S. presidential election. Voting results for the 2008 U.S. Presidential election were obtained from CNN [3]. After the Databionic ESOM software finished clustering counties in a state, some clusters of counties with similar demographics were then analyzed for how they voted in the 2008 U.S. Presidential election.

3.3 Clustering by k-Means

Clustering by k-means was chosen so that it could be compared to ESOM clustering. For a thorough explanation of k-means clustering, please see [8]. The k-means clustering method was used to divide the counties into $k = 30$ clusters. The choice of $k = 30$ clusters was used to make the cluster sizes for ESOMs and for k-means approximately the same. The k-means clustering was performed by the Lloyd-Forgy algorithm as implemented by the Kmeans function in the amap package of the statistics software R [19]. The initial means (or weight vectors) were chosen at random from the data, and a maximum of 100 iterations were allowed. As with the ESOMs, for k-means clustering the variables were normalized using a z-transform and Pearson distance was used as the metric.

3.4   Comparison of Clusterings Using Variation of Information

To measure the dissimilarity between ESOM and k-means clusterings, the variation of information distance was calculated between the ESOM clusterings and the k-means clusterings ($k = 30$). For more information on the variation of information distance metric, please see [15].

4   Results

This section presents the results obtained clustering the census data by ESOMs and k-means clustering. Also, clustering results for ESOMs and k-means are compared using the variation of information distance metric.

4.1   Results for ESOMs

The ESOMs for clustering counties in Florida, Pennsylvania, and Ohio are in Figures 1-3. These ESOMs use U-matrix maps to visualize the correlation between counties in these high-dimensional data sets, thereby creating a topographical map that shows mountain ranges for cluster boundaries, which are indicated by darker shading [23]. These ESOMs are toroidal maps, i.e., the top and bottom edges are identified, and the left and right edges are also identified. This means, for example, that Citrus, Charlotte, Sarasota, and Indian River counties belong to the same cluster in Figure 1, and that Ashland, Mercer, and Tuscarawas counties belong to the same cluster in Figure 3.

The ESOM for Florida in Figure 1 shows that, in general, counties in Florida clustered together by demographics tended to vote for the same presidential candidate in 2008. One notable exception to this was the cluster with Duval, Hillsborough, and Orange counties, which contain the major cities of Jacksonville, Tampa Bay, and Orlando, respectively. How the counties in this cluster voted in the 2008 U.S. presidential election is given in Table 2. The margins for victory in two of these three counties were very narrow. These three counties have very similar demographics since they're in the same cluster in the ESOM. Since political strategies often target particular demographics, strategies that worked in one county for a particular demographic should also work in another county that has similar demographics. That is, political parties could use their winning strategies from one of these three counties in the 2008 election to help them win in the other counties in this cluster in future elections.

The ESOM in Figure 2 shows that the counties in Pennsylvania clustered together by demographics also tended to vote for the same presidential candidate in 2008. The cluster containing Berks, Lancaster, Lehigh, and York counties contains four geographically contiguous counties with 13.5% of the population of the state. Two of these four counties voted for Obama in 2008 while the other two voted for McCain, as shown in Table 2. Since the 2012 U.S. presidential election is projected to be very close in Pennsylvania [1], political parties should consider applying winning strategies from one county in this cluster to all of the counties
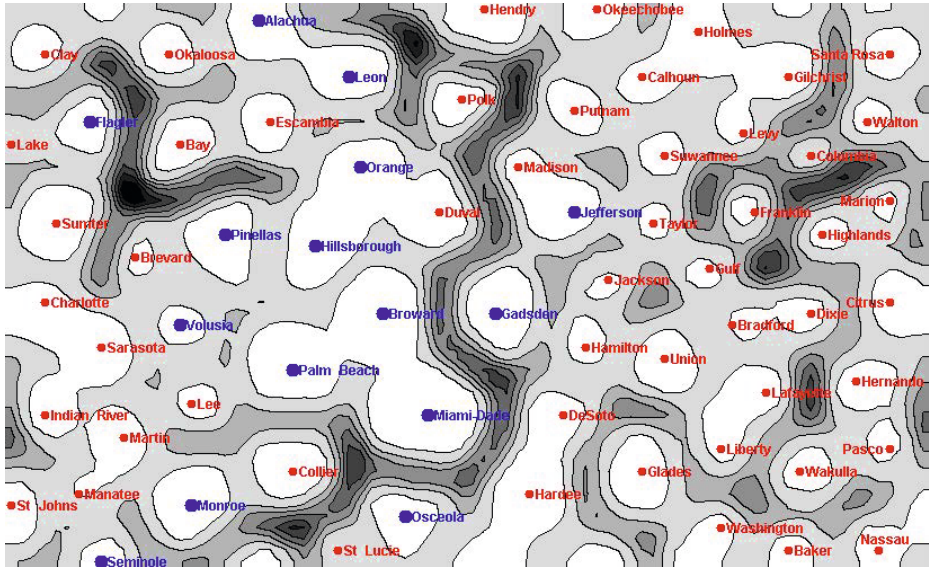
Fig. 1. ESOM clustering of Florida counties by year 2010 U.S. Census data on a toroidal map. Larger blue dots indicate counties that voted for Obama in 2008, while smaller red dots indicate counties that voted for McCain in 2008.

Table 2. (Left) One cluster of counties from the Florida ESOM and how the counties in this cluster voted in the 2008 U.S. presidential election. (Right) One cluster of counties from the Pennsylvania ESOM and how the counties in this cluster voted in the 2008 U.S. presidential election.

Florida Voting (2008)			Pennsylvania Voting (2008)		
County	McCain	Obama	County	McCain	Obama
Duval	51%	49%	Berks	45%	54%
Hillsborough	46%	53%	Lancaster	55%	44%
Orange	41%	59%	Lehigh	42%	57%
			York	56%	43%

in this cluster to try to swing the vote in their favor. Also from the ESOM in Figure 2, there is a cluster that contains Cumberland and Chester counties, which are near the major cities of Harrisburg and Philadelphia, respectively, voted differently in 2008, and were decided by narrow margins. Since these two counties are closest on the ESOM, they have very similar demographics, which suggests that demographically based political strategies that were successful in one county may have a positive effect in the other county.

As was the case with Florida and Pennsylvania, the ESOM in Figure 3 shows that counties in Ohio clustered together by demographics also tended to vote for
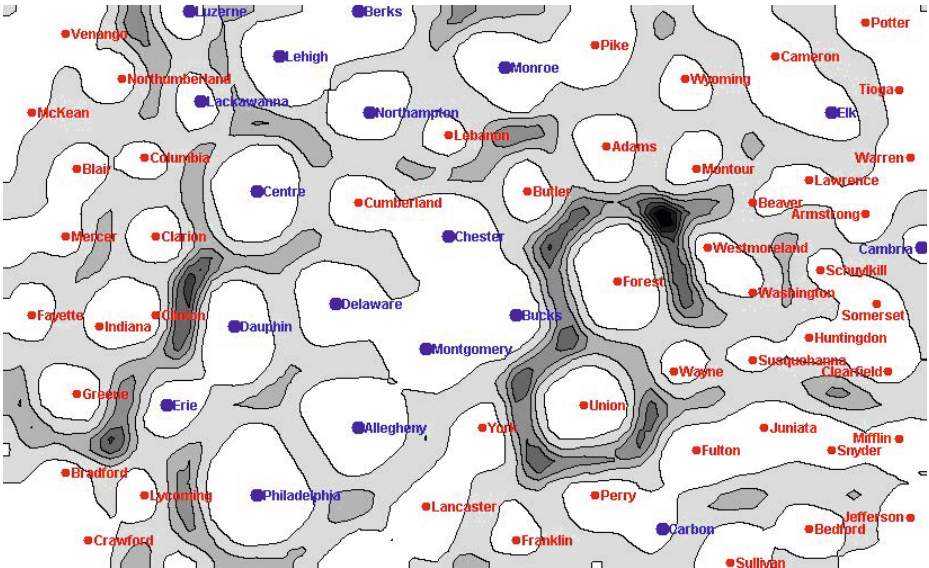
Fig. 2. ESOM clustering of Pennsylvania counties by year 2010 U.S. Census data on a toroidal map. Larger blue dots indicate counties that voted for Obama in 2008, while smaller red dots indicate counties that voted for McCain in 2008.

Fig. 3. ESOM clustering of Ohio counties by year 2010 U.S. Census data on a toroidal map. Larger blue dots indicate counties that voted for Obama in 2008, while smaller red dots indicate counties that voted for McCain in 2008.
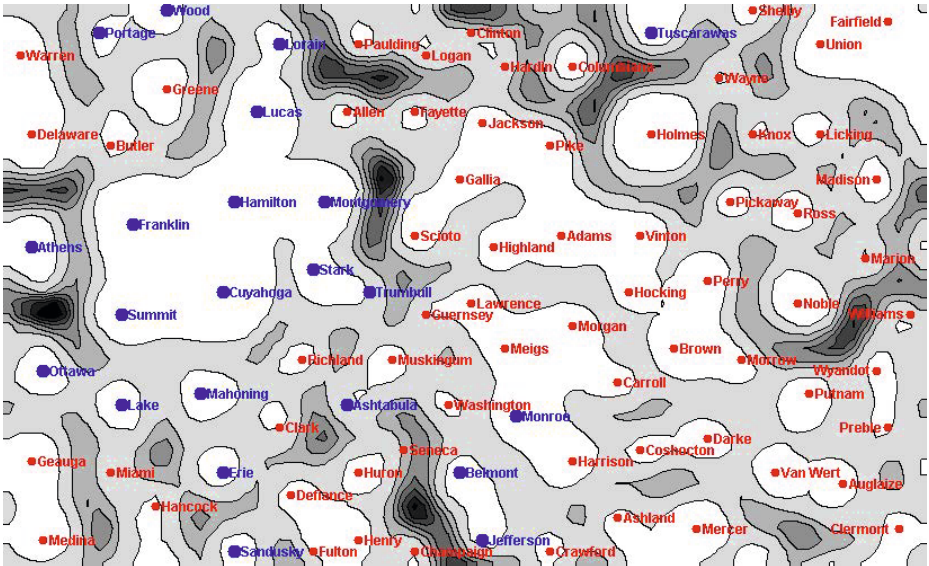
the same presidential candidate in 2008. The cluster containing Ashland, Mercer, and Tuscarawas counties contains one county, Tuscarawas, that voted for Obama over McCain by a margin of 50% to 48%, and two other counties that voted for McCain by substantial margins. This suggests that if the Republican party had used strategies from Ashland and Mercer counties in Tuscarawas county, McCain may have been able to win Tuscarawas county in 2008. A similar argument could be made that political strategies from Meigs or Harrison counties should be applied to Monroe county.

4.2   Results for k-Means Clustering

The results obtained using k-means clustering to create $k = 30$ clusters for the counties in Florida, Pennsylvania, and Ohio are in Tables 3-5. For the these three states, the number of ESOM clusters which have mixed voting results (some counties that voted for Obama and other counties that voted for McCain) are 2 for Florida, 2 for Pennsylvania, and 3 for Ohio. In contrast, the number of k-means clusters ($k = 30$) with such mixed voting results are 9, 4, and 7, respectively. Thus, assuming there is a very close relationship between demographics and presidential voting, the fact that the ESOMs have fewer clusters with mixed voting results may indicate that ESOMs perform better at constructing homogeneous clusters than k-means clustering. Detailed qualitative analysis of the k-means clusterings also reveals results such as that Okaloosa and Miami-Dade counties are in the same cluster, which is somewhat surprising since Okaloosa is considered to be one of the most conservative (Republican) counties in Florida while Miami-Dade is one of the most liberal (Democrat). This may be evidence that the choice of $k = 30$ for k-means clustering of the counties in Florida was perhaps not optimal.

Table 3. Clustering of counties in Florida by year 2010 U.S. Census data via k-means (with $k = 30$) together with how each county voted in the 2008 U.S. presidential election (D = Democrat = Obama, R = Republican = McCain).

Cluster size	Clusters
6	{(R) Franklin, (R) Hernando, (D) Jefferson, (R) Lake, (R) St. Johns, (R) Santa Rosa}
5	{(R) Hardee, (D) Hillsborough, (R) Sarasota, (R) Wakulla, (R) Washington}
4	{(R) Collier, (R) Dixie, (D) Flagler, (R) Indian River}
3	{(R) Holmes, (D) Miami-Dade, (R) Okaloosa}, {(R) Calhoun, (R) Gulf, (R) St. Lucie}, {(R) Brevard, (R) Citrus, (R) Glades}, {(R) Lafayette, (R) Liberty, (R) Martin}, {(R) Duval, (R) Sumter, (R) Taylor}, {(R) Bradford, (R) Gilchrist, (R) Jackson}, {(R) Hamilton, (D) Leon, (R) Nassau}, {(R) Clay, (R) Hendry, (R) Madison}
2	{(R) Manatee, (R) Walton}, {(R) Marion, (D) Pinellas}, {(D) Monroe, (R) Polk}, {(R) Charlotte, (R) Baker}, {(D) Orange, (D) Volusia}, {(D) Palm Beach, (D) Seminole}, {(D) Gadsen, (R) Pasco}, {(R) Highlands, (R) Union}, {(R) Lee, (D) Osceola}
1	{(R) Columbia}, {(R) Bay}, {(D) Broward}, {(R) Okeechobee}, {(R) Putnam}, {(R) Suwanee}, {(R) Escambia}, {(R) DeSoto}, {(R) Alachua}, {(R) Levy}

Table 4. Clustering of counties in Pennsylvania by year 2010 U.S. Census data via k-means (with $k = 30$) together with how each county voted in the 2008 U.S. presidential election (D = Democrat = Obama, R = Republican = McCain).

Cluster size	Clusters
11	{(R) Bedford, (R) Bradford, (R) Clearfield, (R) Crawford, (R) Jefferson, (R) Lycoming, (R) McKean, (R) Potter, (R) Tioga, (R) Venango, (R) Warren}
5	{(R) Beaver, (R) Blair, (D) Cambria, (R) Lawrence, (R) Mercer}, {(R) Clarion, (R) Clinton, (R) Columbia, (R) Greene, (R) Indiana}, {(R) Armstrong, (R) Huntingdon, (R) Schuylkill, (R) Somerset, (R) Susquehanna}
3	{(R) Cameron, (D) Elk, (R) Montour}
2	{(D) Monroe, (R) Pike}, {(R) Sullivan, (R) Wayne}, {(R) Franklin, (R) Lebanon}, {(R) Butler, (R) Washington}, {(R) Adams, (R) Wyoming}, {(D) Lackawanna, (D) Luzerne}, {(R) Forest, (R) Union}, {(D) Bucks, (D) Montgomery}, {(R) Mifflin, (R) Northumberland}, {(D) Carbon, (R) Fulton}, {(D) Dauphin, (D) Erie}, {(R) Juniata, (R) Snyder}, {(D) Allegheny, (D) Philadelphia}
1	{(R) Cumberland}, {(D) Chester}, {(D) Northampton}, {(D) Lehigh}, {(D) Berks}, {(D) Centre}, {(R) Fayette}, {(R) Perry}, {(R) Lancaster}, {(R) York}, {(D) Delaware}, {(R) Westmoreland}

Table 5. Clustering of counties in Ohio by year 2010 U.S. Census data via k-means (with $k = 30$) together with how each county voted in the 2008 U.S. presidential election (D = Democrat = Obama, R = Republican = McCain).

Cluster size	Clusters
7	{(R) Carroll, (R) Guernsey, (R) Highland, (R) Lawrence, (R) Meigs, (R) Morgan, (R) Washington}, {(R) Adams, (R) Fayette, (R) Gallia, (R) Jackson, (R) Pike, (R) Scioto, (R) Vinton}
6	{(D) Belmont, (R) Coshocton, (R) Crawford, (R) Harrison, (D) Jefferson, (D) Monroe}, {(R) Brown, (R) Darke, (R) Hocking, (R) Knox, (R) Morrow, (R) Perry}
5	{(D) Ottawa, (R) Preble, (R) Shelby, (R) Williams, (R) Wyandot}, {(R) Greene, (R) Hancock, (R) Miami, (D) Portage, (D) Wood}, {(R) Defiance, (R) Fulton, (R) Henry, (R) Huron, (D) Sandusky}, {(D) Cuyahoga, (D) Hamilton, (D) Montgomery, (D) Stark, (D) Summit}
4	{(R) Ashland, (R) Auglaize, (R) Mercer, (R) Putnam}, {(R) Columbiana; (R) Marion, (R) Noble, (R) Pickaway}
3	{(R) Clark, (D) Erie, (D) Mahoning}, {(R) Paulding, (R) Seneca, (R) Van Wert}, {(R) Delaware, (R) Fairfield, (R) Warren}, {(R) Allen, (R) Richland, (D) Trumbull}
2	{(R) Geauga, (R) Medina}, {(R) Madison, (R) Ross}, {(D) Lorain, (D) Lucas}, {(R) Ashtabula, (D) Muskingum}, {(R) Clinton, (R) Hardin}, {(R) Butler, (D) Franklin}
1	{(R) Wayne}, {(R) Holmes}, {(D) Athens}, {(R) Champaign}, {(R) Logan}, {(D) Tuscarawas}, {(R) Clermont}, {(R) Union}, {(D) Lake}, {(R) Licking}

4.3   Comparison of Clusterings Using Variation of Information

To compare the ESOM and k-means clusterings for each state, variation of information distance (with a base 2 logarithm) was used [15]. The variation of information results are shown in Table 6, and they were produced using a Matlab / Octave function written by the first author. Since ESOM clustering used a U-matrix, its clustering results show somewhat continuous variation between clusters. In contrast, the variation between k-means clusters is discrete. To make the ESOM clustering results discrete for the purpose of using variation of

Table 6. Variation of information (VI) distances between ESOM clustering and k-means clustering ($k = 30$)

	Florida	Pennsylvania	Ohio
Actual VI distance	1.9856	1.6366	1.6585
Max possible VI distance	$\log_2(67)$	$\log_2(67)$	$\log_2(88)$
(Actual VI) / (Max possible VI)	0.32733	0.26979	0.25676

information to compare it with k-means clustering, counties in an ESOM were considered to be in distinct clusters whenever a level curve on the topographical map would have to be crossed, and in the same cluster otherwise. Since none of the variation of information distances in Table 6 are close to zero, the ESOM and k-means clusterings are measurably different. The results in Table 6 support the claim made by Ultsch and Moerchen [23] that ESOM clusterings are different from k-means clusterings, which are believed to be very similar to ordinary SOM clusterings.

5   Discussion

The article by Nate Silver [20] emphasizes that every individual is a member of many demographic categories and "the voting tendencies associated with those categories often point in different, or even conflicting, directions." Despite this, political parties have tried to target individual demographic categories instead of incorporating strategies that address several demographic categories simultaneously. Using an ESOM to cluster geographic regions by a broadly defined set of demographic data could help political parties identify possibly non-contiguous geographic regions that could benefit from demographically targeted political strategies. Additionally, using an ESOM to cluster groups of people across many demographic categories at once, and then pairing this information with voting tendencies could help produce a clearer picture of how demographics are related to voting tendencies. This information could play a vital and pivotal role for political parties trying to win elections in very close races.

Future directions for research in this area may include sensitivity analysis of ESOMs obtained by looking at how an ESOM changes when one of its input variables changes, comparing the results of ESOMs to other exploratory clustering methods by using variation of information or some other method of comparing clusterings, or using ESOMs in a similar manner on a smaller, more narrowly defined, set of demographic variables to try to produce ESOMs having even fewer clusters with non-homogeneous voting outcomes.

References

1. 2012 Presidential Election Interactive Map and History of the Electoral College, `http://www.270towin.com/` (retrieved May 5, 2012)
2. Ansolabehere, S., Persily, N., Stewart, C.: Race, Region, and Vote Choice in the 2008 Election: Implications for the Future of the Voting Rights Act. Harvard Law Review 123 (2010); Columbia Public Law Research Paper No. 09-211; MIT Political Science Department Research Paper No. 2011-1. Available at SSRN: `http://ssrn.com/abstract=1462363` (retrieved May 1, 2012)
3. CNN. County Results - Election Center 2008 - Elections & Politics from CNN.com, `http://www.cnn.com/ELECTION/2008/results/county/` (retrieved May 1, 2012)
4. Cooper, C., Burns, A.: Kohonen Self-Organizing Feature Maps as a Means to Benchmark College and University Websites. Journal of Science Education and Technology 16(3), 203–211 (2007)
5. Frey, W.: Battling Battlegrounds. American Demographics (September 24–26, 2004)
6. Gelman, A., Kenworthy, L., Su, Y.: Income Inequality and Partisan Voting in the United States. Social Science Quarterly, Special Issue: Inequality and Poverty: American and International Perspectives 91(5), 1203–1219 (2010)
7. Gimpel, J., Dyck, J., Shaw, D.: Registrants, Voters, and Turnout Variability Across Neighborhoods. Political Behavior 26(4), 343–375 (2004)
8. Hartigan, J.: Clustering Algorithms, pp. 1–351. Wiley, New York (1975)
9. Kaski, S., Kohonen, T.: Exploratory Data Analysis by the Self-Organizing Map: Structures of Welfare and Poverty in the World. In: Refenes, A., Abu-Mostafa, Y., Moody, J., Weigend, A. (eds.) Neural Networks in Financial Engineering, pp. 498–507. World Scientific, Singapore (1996)
10. Kohonen, T.: Self-Organizing Maps, 3rd edn., pp. 1–521. Springer, Berlin (2000)
11. Kohonen, T.: The Self-Organizing Map. Proceedings of the IEEE 78(9), 1464–1480 (1990)
12. Laaksonen, J., Honkela, T. (eds.): WSOM 2011. LNCS, vol. 6731, pp. 1–380. Springer, Heidelberg (2011)
13. Lesthaeghe, R., Niedert, L.: US Presidential Elections and the Spatial Pattern of the American Second Demographic Transition. Population and Development Review 35(2), 391–400 (2009)
14. Lopez, M.: Dissecting the 2008 Electorate: Most Diverse in U.S. History - Pew Research Center, `http://pewresearch.org/pubs/1209/racial-ethnic-voters-presidential-election` (retrieved May 1, 2012)
15. Meilă, M.: Comparing Clusterings by the Variation of Information. In: Schölkopf, B., Warmuth, M.K. (eds.) COLT/Kernel 2003. LNCS (LNAI), vol. 2777, pp. 173–187. Springer, Heidelberg (2003)
16. Neme, A., Hernández, S., Neme, O.: An Electoral Preferences Model Based on Self-Organizing Maps. Journal of Computational Science 2, 345–352 (2011)
17. Neme, A., Hernández, S., Neme, O.: Self Organizing Maps as Models of Social Processes: The Case of Electoral Preferences. In: Laaksonen, J., Honkela, T. (eds.) WSOM 2011. LNCS, vol. 6731, pp. 51–60. Springer, Heidelberg (2011)
18. Niemelä, P., Honkela, T.: Analysis of Parliamentary Election Results and Socio-Economic Situation Using Self-Organizing Map. In: Príncipe, J.C., Miikkulainen, R. (eds.) WSOM 2009. LNCS, vol. 5629, pp. 209–218. Springer, Heidelberg (2009)
19. The R project for statistical computing, `http://www.r-project.org/` (retrieved June 20, 2012)

20. Silver, N.: In Politics, Demographics Are Not Destiny - NYTimes.com, `http://fivethirtyeight.blogs.nytimes.com/2011/03/01/in-politics-demographics-are-not-destiny/` (retrieved May 1, 2012)
21. Trosset, M.: Representing Clusters: K-Means Clustering, Self-Organizing Maps, and Multidimensional Scaling, Technical Report 08-03, Department of Statistics, Indiana University, Bloomington, IN (2008)
22. Tuia, D., Kaiser, C., Da Cunha, A., Kanevski, M.: Socio-economic Data Analysis with Scan Statistics and Self-organizing Maps. In: Gervasi, O., Murgante, B., Laganà, A., Taniar, D., Mun, Y., Gavrilova, M.L. (eds.) ICCSA 2008, Part I. LNCS, vol. 5072, pp. 52–64. Springer, Heidelberg (2008)
23. Ultsch, A., Moerchen, F.: ESOM-Maps: tools for clustering, visualization, and classification with Emergent SOM, Technical Report Dept. of Mathematics and Computer Science, University of Marburg, Germany, No. 46 (2005)
24. United States Census Bureau. Download QuickFacts from the US Census Bureau, `http://quickfacts.census.gov/qfd/download_data.html` (retrieved May 1, 2012)
25. United States Census Bureau. Dictionary of census data, `http://quickfacts.census.gov/qfd/download/DataDict.txt` (retrieved May 1, 2012)