Dynamic Graphics: An Interactive Analysis Of What Attaches People To Their Communities

Jessica M. Orth*

Abstract
In this research, we will investigate several different approaches and methods to displaying multivariate data. Emphasis will be placed on end-user-customization tools and flexibility in dynamic and interactive displays. Specifically, we will highlight the use of motion charts using Markus Gesmann's `googleVis` package in R. We will demonstrate the visualization of time-series data and also the results of Multidimensional Scaling and Principal Component Analysis using this tool. The goals of these displays are ease of usability and interpretation, dynamic customization options, and the ability to display multivariate data in a meaningful way. We will use data collected from the Knight Foundation and Gallup during the years 2008-2010 to illustrate the attachment of people to their communities in a new and innovative way.

Key Words: Dynamic graphics; Interactive visualizations; Multivariate statistical analysis

1. Approach

Displaying multivariate data can be achieved in many ways through a variety of tools. Here we aim to emphasize the use of motion charts for displaying the trend analysis of time-dependent Principal Component Analysis and Multidimensional Scaling. It is well known that these methods are used as data reduction and data mining techniques in the analysis of multivariate data, but what happens when we introduce a time variable to these results? As will be seen, motion charts provide the tool to seamlessly merge these results throughout time and allow for dynamic and interactive interpretations of what attaches people to their communities.

We analyze the results of 43,000 people from 26 communities across America using the index variables collected from the 'Soul of the Community' survey conducted by the Knight Foundation and Gallop. The cities are shown in the following figure.

*Department of Statistics and Actuarial Science, The University of Iowa, 241 Schaeffer Hall Iowa City, IA 52242-1409

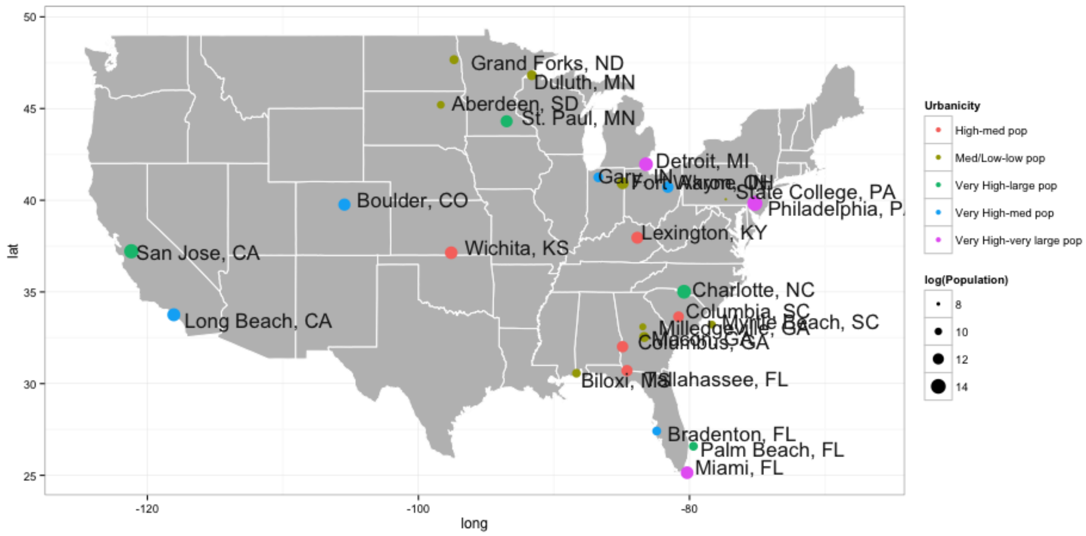


Figure 1: Locations of communities surveyed

The index variables analyzed include the specific attributes that make people live where they live such as: attachment, loyalty, passion, basic services, leadership, education, safety, aesthetics, economy, social offerings, community offerings, civic involvement, openness, social capital, and community domains. Our analysis looks at four different summary statistics: means, standard deviations, the proportion of high index variables, and z-scores.

Means, standard deviations, and proportions are calculated for cities based on the categorical responses on scales of 0-3 and 0-5. The z-scores serve as an index themselves since they are calculated for each index variable by city in relation to the collection of responses for the index variables. This provides information on each city's score for the original index variables: negative z-scores imply a lower score for the index variable and positive z-scores indicate a higher score for that city, relative to the overall score of the original index variable.

Data reduction and data mining techniques are applied to these summary statistics and the results of Principal Component Analysis and Multidimensional Scaling are displayed dynamically in interactive motion charts. To view the motion charts dynamically, see `http://mnstats.morris.umn.edu/JSM2013.html`. In Section 2 we investigate how the index variables are related to each other by identifying the dynamic drivers that affect community attachment. In Section 3.1 we examine the relationships between the cities and consider clusters of similar cities as well as explore the movement of dynamic cities throughout the three years surveyed. We also use the means for the index variables to analyze average hierarchical clustering of the cities through dendrograms in Section 3.2. Conclusions and future research are addressed in Section 4.

Here is a summary of our approach.

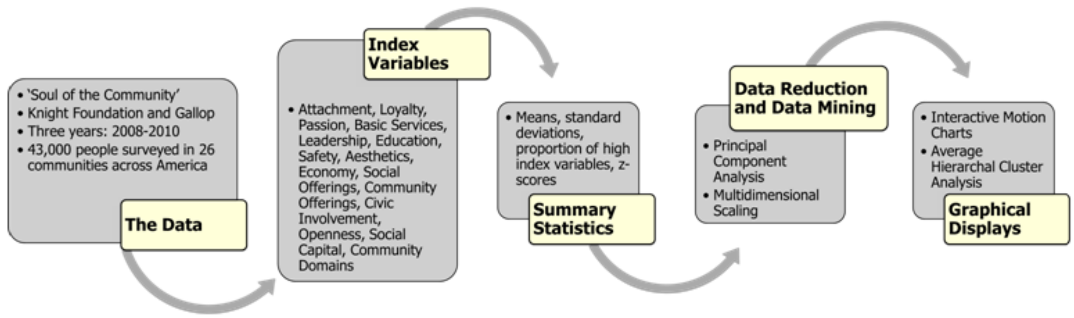


Figure 2: Diagram of our approach

2. Key Drivers and Relationships Between Them (PCA)

Principal Component Analysis seeks to create uncorrelated linear combinations of the index variables that explain a maximal amount of variation in the data. In traditional Principal Component Analysis, the set of linear combinations of the variables with the greatest variance is used as the first principal component and has the following form:

$$y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p \tag{1}$$

with the constraint that $\mathbf{a_1'a_1} = 1$. Subsequent principal components are constructed similarly, with the general form

$$y_k = a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kp}x_p = \mathbf{a_k'x} \tag{2}$$

with constraints $\mathbf{a_k'a_k} = 1$ and $\mathbf{a_k'a_i} = 0$ for $(i < k)$.

We construct time-dependent principal components, taking into account the three years surveyed, and expressed as

$$y_k(t) = a_{k1}(t)x_1(t) + a_{k2}(t)x_2(t) + \ldots + a_{kp}(t)x_p(t) = \mathbf{a_k(t)'x(t)} \tag{3}$$

with constraints $\mathbf{a_k(t)'a_k(t)} = 1$ and $\mathbf{a_k(t)'a_i(t)} = 0$ for $(i < k)$.

We investigate the first two principal components obtained from the summary statistics. The following tables show the loadings from the analysis. Note the changes in the loadings as we move from Dimension 1 to Dimension 2 throughout time. For the means, the loadings are positive and mostly close to 1 in Dimension 1, and they change to mostly negative and closer to 0 in Dimension 2. We see a contrast between social capital and safety vs. all other index variables in Dimension 1 for the standard deviations in 2008, while the contrast in 2009 and 2010 is only between safety and all other index variables for Dimension 1. The loadings for proportions show a contrast of safety and aesthetics vs. all other index variables in 2008 for Dimension 1, while in 2009 the contrast is between social capital vs. all other index variables. In 2010 we obtain the overall effect of attachment without contrasts. Leadership in Dimension 2 for proportions changes from positive to negative to postive throughout the three years, while economy and openness change from negative to positive to negative. Social offerings are positive in 2008 and negative in 2009-2010, while community offerings have the opposite effect.

The first dimension of the principal component analysis serves as an index for the overall drivers of attachment for each summary statistic, while the second dimension shows a contrast between economic growth and emotional bond. We are able to explain approximately 55-75% of the variation in the data with the first two dimensions.

	Dimension 1			Dimension 2		
Index Variable	2008	2009	2010	2008	2009	2010
Loyalty	0.920	0.898	0.918	-0.224	-0.335	-0.285
Passion	0.890	0.887	0.894	-0.347	-0.394	-0.382
Community Attachment	0.916	0.897	0.910	-0.297	-0.370	-0.340
Basic Services	0.286	0.365	0.437	0.502	0.349	0.120
Leadership	0.744	0.777	0.790	0.264	0.248	0.391
Education	0.729	0.772	0.837	0.464	0.380	0.311
Safety	0.499	0.625	0.704	0.653	0.616	0.533
Aesthetics	0.570	0.676	0.748	-0.147	-0.210	-0.283
Economy	0.660	0.697	0.819	-0.026	0.112	0.315
Social Offerings	0.729	0.791	0.806	-0.405	-0.392	-0.386
Community Offerings	0.947	0.966	0.980	0.243	0.175	0.130
Involvement	0.568	0.592	0.413	-0.042	0.240	0.035
Openness	0.682	0.744	0.756	-0.542	-0.451	-0.490
Social Capital	0.496	0.421	0.489	0.706	0.737	0.744
Domains	0.967	0.972	0.982	-0.007	0.074	0.012

Table 1: Loadings for PCA on means for survey years 2008-2010

	Dimension 1			Dimension 2		
Index Variable	2008	2009	2010	2008	2009	2010
Loyalty	0.412	0.540	0.502	0.835	0.788	0.815
Passion	0.417	0.500	0.460	0.845	0.809	0.801
Community Attachment	0.466	0.548	0.502	0.854	0.802	0.815
Basic Services	0.601	0.653	0.610	-0.009	-0.009	0.281
Leadership	0.672	0.569	0.437	-0.509	-0.553	-0.697
Education	0.580	0.590	0.533	-0.111	0.128	0.155
Safety	-0.206	-0.170	-0.328	0.188	-0.182	0.045
Aesthetics	0.708	0.656	0.643	0.299	0.356	0.432
Economy	0.308	0.410	0.244	-0.653	-0.674	-0.830
Social Offerings	0.700	0.654	0.724	-0.422	-0.484	-0.380
Community Offerings	0.949	0.912	0.868	-0.141	-0.280	-0.200
Involvement	0.182	0.600	0.675	-0.250	-0.343	-0.225
Openness	0.823	0.733	0.758	-0.027	-0.247	-0.300
Social Capital	-0.023	0.406	0.458	0.389	0.524	-0.004
Domains	0.637	0.838	0.680	-0.233	-0.317	-0.539

Table 2: Loadings for PCA on standard deviations for survey years 2008-2010

	Dimension 1			Dimension 2		
Index Variable	2008	2009	2010	2008	2009	2010
Loyalty	0.953	0.886	0.894	-0.045	-0.293	-0.259
Passion	0.724	0.756	0.829	0.319	0.259	0.147
Community Attachment	0.947	0.890	0.897	-0.040	-0.289	-0.261
Basic Services	0.211	0.457	0.377	-0.421	-0.417	-0.598
Leadership	0.733	0.777	0.853	0.028	-0.067	0.057
Education	0.447	0.617	0.724	0.566	0.464	0.394
Safety	-0.030	0.295	0.347	0.695	0.673	0.768
Aesthetics	-0.085	0.058	0.319	0.673	0.775	0.605
Economy	0.723	0.881	0.637	-0.173	0.051	-0.196
Social Offerings	0.850	0.888	0.860	0.016	-0.131	-0.115
Community Offerings	0.293	0.695	0.156	-0.480	-0.025	0.279
Involvement	0.186	0.274	0.551	0.838	0.587	0.279
Openness	0.623	0.465	0.315	-0.299	0.121	-0.502
Social Capital	0.051	-0.357	0.165	0.517	0.262	0.575

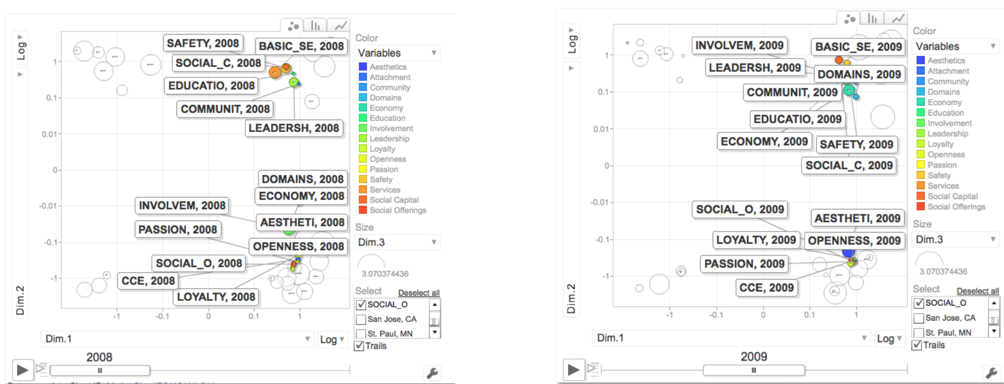
Table 3: Loadings for PCA on proportions for survey years 2008-2010

We are interested in the change in the relationship of the index variables from year to year, and the motion charts clearly show the dynamic drivers of attachment for each summary statistic.

One of the many beauties of motion charts is the capability to put the analysis into the hands of the user. Rather than limit a client with one simple graphical display, motion charts allow for customizable analyses to suit the interests of multiple users. All one has to do is change the axes, or modify the color variable or size variable, to create a unique analysis that is more informative than a single display.

While social offerings, openness, and aesthetics are found to be the leading drivers of community attachment by the Knight Foundation, we are able to examine the relationship between these and the other index variables easily with the motion charts. The first principal component is plotted on the x-axis and the second principal component on the y-axis, with the color specified by the index variables. By viewing the display on the log-log scale, which only changes the axes scale, we are able to observe two distinct clusters in each of the summary statistics: overall drivers of attachment and emotional bond.

The following figures and table summarize these results.



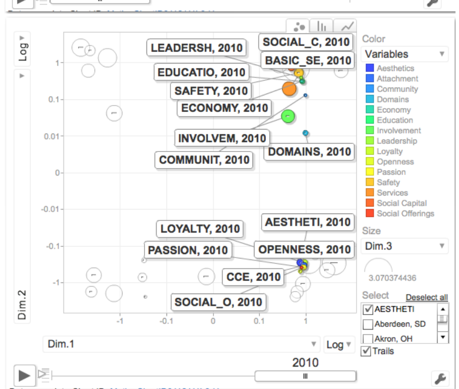


Figure 3: PCA results of the means for 2008-2010. Dimension 1 is on the x-axis and Dimension 2 is on the y-axis. The data are shown on the log-log scale, with color specified by the index variables.

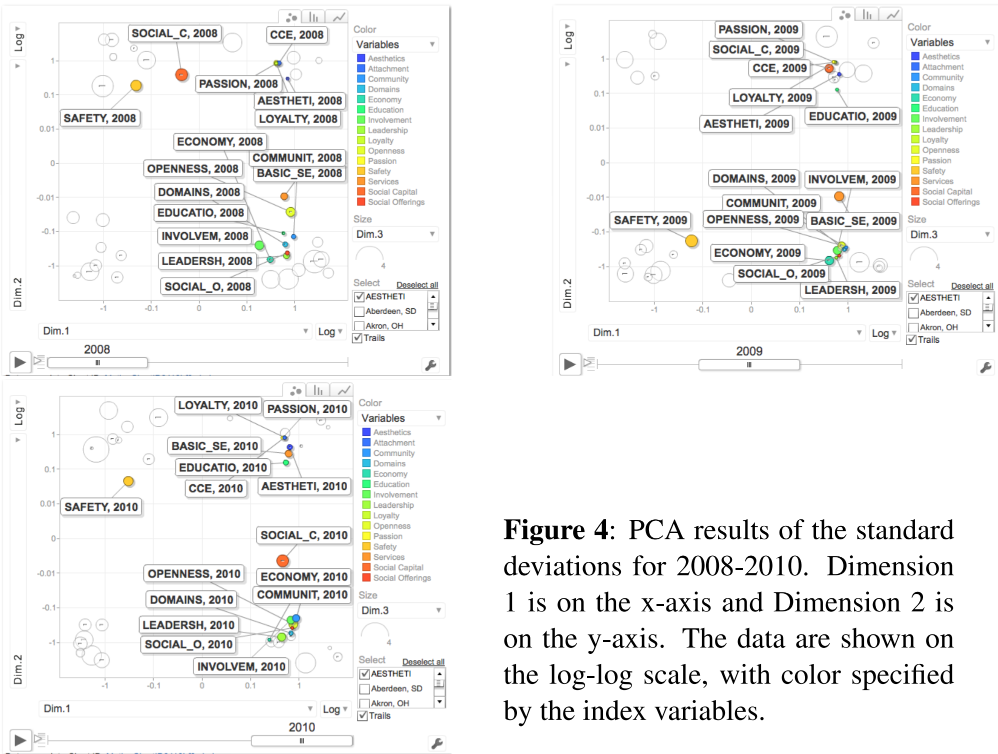


Figure 4: PCA results of the standard deviations for 2008-2010. Dimension 1 is on the x-axis and Dimension 2 is on the y-axis. The data are shown on the log-log scale, with color specified by the index variables.

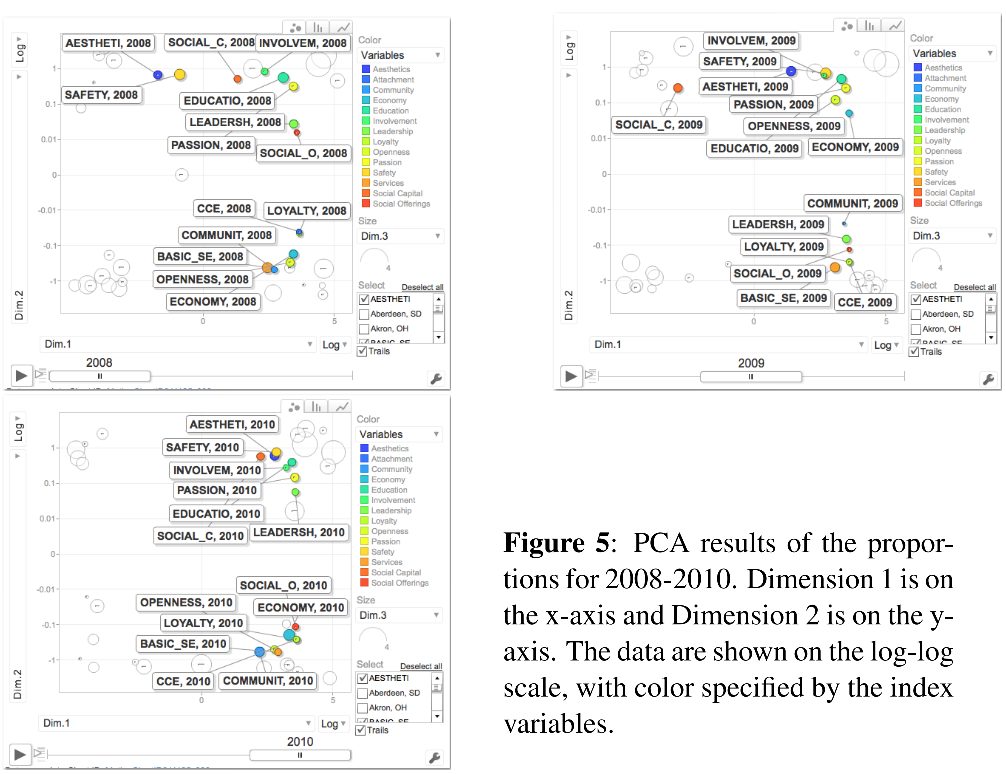


Figure 5: PCA results of the proportions for 2008-2010. Dimension 1 is on the x-axis and Dimension 2 is on the y-axis. The data are shown on the log-log scale, with color specified by the index variables.

	Means	Standard Deviations	Proportions
Dimension 1	Overall drivers for attachment	Personal Assurance vs. Overall drivers for attachment	Personal Assurance vs. Overall drivers for attachment
Percentage of Variation Explained	2008: 54 2009: 58 2010: 62	2008: 32 2009: 38 2010: 34	2008: 35 2009: 42 2010: 39
Dimension 2	Economic Growth vs. Emotional Bond	Personal Assurance and Pride vs. Economic Growth	Emotional Bond vs. Economic Growth
Percentage of Variation Explained	2008: 15 2009: 14 2010: 13	2008: 23 2009: 25 2010: 27	2008: 20 2009: 15 2010: 20
Dynamic Drivers	Involvement, Economy, Domains	Safety, Social Capital, Education, Basic Services	Safety, Aesthetics, Social Capital, Leadership, Social Offering, Openness, Economy

Table 4: Dynamic drivers and percentage of variation explained by Principal Component Analysis

3. Differences Between Communities

3.1 Multidimensional Scaling

The goal of Multidimensional Scaling is to provide a visual representation of the pattern of similarities and differences among the cities. By calculating the Euclidean distance between the points, MDS maps the cities based on proximity matrices. Let $d_{ij}(t)$ be the distance between coordinates $\mathbf{x_i(t)}$ and $\mathbf{x_j(t)}$ for different time periods. Then the Euclidean distance is calculated as

$$d_{ij}(t) = \sqrt{\sum_{k=1}^{d} (x_{ik}(t) - x_{jk}(t))^2} \tag{4}$$

We use the index variables to determine the relationships between the cities. Cities estimated to be very similar to each other in these characteristics are placed close to each other on the map, and those estimated to be very different from each other are placed far away from each other on the map. The following tables show the loadings from the analysis.

	Dimension 1			Dimension 2		
Index Variable	2008	2009	2010	2008	2009	2010
Aberdeen, SD	0.544	0.560	0.467	0.195	-0.467	-0.489
Akron, OH	-0.500	-0.499	-0.537	-0.063	-0.020	-0.040
Biloxi, MS	0.225	0.182	0.297	0.528	-0.408	0.289
Boulder, CO	0.621	0.961	0.968	-0.460	0.036	-0.125
Bradenton, FL	0.230	0.648	0.568	-0.284	0.345	0.227
Charlotte, NC	0.024	0.116	0.041	0.057	0.0534	0.097
Columbia, SC	0.117	0.019	-0.039	0.151	0.141	0.078
Columbus, GA	0.182	0.057	0.298	0.167	0.072	0.134
Detroit, MI	-1.293	-1.420	-1.426	-0.109	-0.246	-0.088
Duluth, MN	0.420	0.431	0.397	-0.306	-0.047	-0.187
Fort Wayne, IN	-0.199	-0.143	0.245	0.079	0.091	0.073
Gary, IN	-1.636	-2.304	-2.111	0.024	-0.068	-0.117
Grand Forks, ND	0.685	0.842	0.810	0.261	-0.614	-0.548
Lexington, KY	0.369	0.496	0.330	0.019	0.077	-0.003
Long Beach, CA	0.239	0.275	0.380	0.060	0.186	0.217
Macon, GA	-0.477	-1.205	-1.230	0.104	-0.020	0.046
Miami, FL	-0.866	-0.879	-0.858	-0.303	0.376	0.236
Milledgeville, GA	-0.133	0.0213	-0.746	0.141	0.102	0.079
Myrtle Beach, SC	0.365	0.412	0.279	0.295	0.343	0.373
Palm Beach, FL	-0.133	0.064	0.101	-0.299	0.469	0.262
Philadelphia, PA	-0.184	-0.185	-0.359	-0.014	-0.132	0.013
San Jose, CA	0.199	0.147	0.142	-0.087	-0.085	-0.091
St. Paul, MN	0.407	0.661	0.611	-0.152	-0.011	-0.081
State College, PA	0.646	0.738	1.051	-0.160	-0.200	-0.382
Tallahassee, FL	0.222	0.369	0.342	-0.159	0.047	-0.021
Wichita, KS	-0.073	-0.365	-0.021	0.314	-0.021	0.049

Table 5: Loadings for MDS on means for survey years 2008-2010

	Dimension 1			Dimension 2		
Index Variable	2008	2009	2010	2008	2009	2010
Aberdeen, SD	0.181	0.127	0.271	-0.073	-0.089	-0.046
Akron, OH	-0.145	-0.108	-0.086	-0.025	-0.021	0.010
Biloxi, MS	-0.118	-0.132	-0.066	-0.205	-0.138	-0.139
Boulder, CO	0.291	0.277	0.201	0.218	0.138	0.198
Bradenton, FL	0.052	0.188	0.165	0.052	0.027	0.015
Charlotte, NC	-0.071	-0.037	-0.030	0.063	-0.009	0.005
Columbia, SC	0.086	0.075	-0.054	-0.028	-0.054	-0.003
Columbus, GA	-0.050	-0.118	-0.081	-0.148	-0.122	-0.192
Detroit, MI	-0.341	-0.239	-0.256	0.067	0.091	0.031
Duluth, MN	0.224	0.069	0.003	0.127	0.135	0.110
Fort Wayne, IN	-0.088	0.106	0.225	-0.041	0.071	0.026
Gary, IN	-0.402	-0.378	-0.464	0.184	0.128	0.112
Grand Forks, ND	0.125	0.197	0.273	-0.115	-0.140	-0.127
Lexington, KY	0.239	0.196	0.100	-0.062	-0.022	-0.012
Long Beach, CA	0.022	0.006	0.093	-0.101	-0.049	-0.035
Macon, GA	-0.286	-0.405	-0.410	-0.140	-0.017	-0.031
Miami, FL	-0.361	-0.292	-0.246	0.202	0.061	0.061
Milledgeville, GA	-0.141	-0.136	-0.364	-0.102	-0.093	-0.065
Myrtle Beach, SC	0.033	-0.018	0.039	-0.075	-0.029	0.000
Palm Beach, FL	-0.187	0.067	-0.046	0.088	0.026	-0.025
Philadelphia, PA	-0.096	0.053	-0.123	0.042	0.020	0.016
San Jose, CA	0.347	0.155	0.245	-0.033	0.057	0.019
St. Paul, MN	0.163	0.203	0.213	0.023	0.061	0.055
State College, PA	0.394	0.206	0.303	0.047	-0.016	0.004
Tallahassee, FL	0.047	0.051	0.064	0.004	-0.013	0.035
Wichita, KS	0.081	-0.114	0.029	0.029	-0.002	-0.022

Table 6: Loadings for MDS on standard deviations for survey years 2008-2010

	Dimension 1			Dimension 2		
Index Variable	2008	2009	2010	2008	2009	2010
Aberdeen, SD	0.034	-0.005	0.092	0.159	0.223	0.188
Akron, OH	0.108	0.104	0.152	-0.079	-0.053	-0.080
Biloxi, MS	0.122	0.044	0.004	0.168	0.150	0.202
Boulder, CO	-0.493	-0.401	-0.493	-0.123	-0.193	-0.198
Bradenton, FL	-0.095	-0.156	-0.138	0.020	0.073	-0.008
Charlotte, NC	0.054	0.034	0.064	-0.021	-0.012	-0.031
Columbia, SC	0.063	0.068	0.096	0.012	0.029	0.027
Columbus, GA	-0.000	0.038	-0.028	0.072	0.102	0.137
Detroit, MI	0.145	0.171	0.239	-0.121	-0.108	-0.095
Duluth, MN	-0.235	-0.175	-0.236	-0.066	-0.124	-0.114
Fort Wayne, IN	0.098	0.162	0.104	-0.022	-0.040	-0.017
Gary, IN	0.220	0.263	0.294	-0.148	-0.130	-0.080
Grand Forks, ND	0.006	-0.105	-0.159	0.255	0.245	0.223
Lexington, KY	-0.022	-0.009	-0.017	0.006	0.039	-0.001
Long Beach, CA	-0.022	0.001	0.001	0.041	0.045	0.018
Macon, GA	0.106	0.169	0.205	-0.038	-0.047	-0.021
Miami, FL	0.083	0.085	0.172	-0.089	-0.104	-0.083
Milledgeville, GA	0.079	0.040	0.099	0.034	0.103	0.058
Myrtle Beach, SC	-0.039	-0.075	0.012	0.113	0.105	0.078
Palm Beach, FL	-0.027	-0.035	-0.103	-0.062	-0.087	-0.044
Philadelphia, PA	0.079	0.091	0.123	-0.051	-0.064	-0.044
San Jose, CA	0.006	0.015	0.045	-0.032	-0.041	-0.068
St. Paul, MN	-0.111	-0.154	-0.164	-0.029	-0.082	-0.117
State College, PA	-0.185	-0.203	-0.396	0.091	0.020	0.094
Tallahassee, FL	-0.108	-0.110	-0.095	-0.064	-0.052	-0.053
Wichita, KS	0.134	0.144	0.126	-0.026	0.002	0.030

Table 7: Loadings for MDS on proportions for survey years 2008-2010

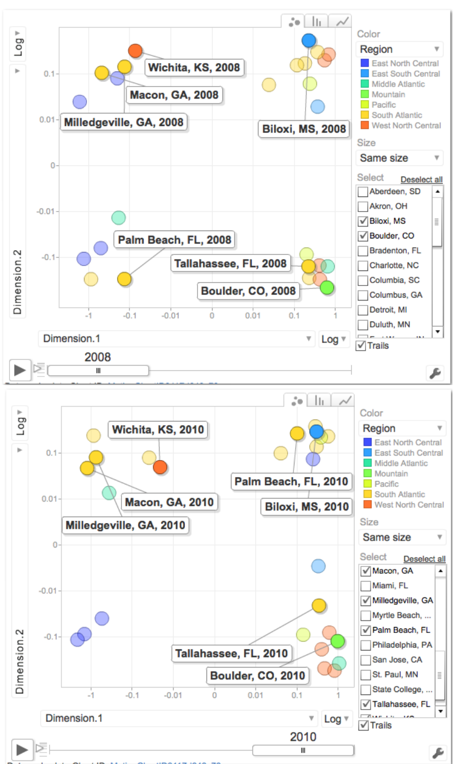
	Dimension 1			Dimension 2		
Index Variable	2008	2009	2010	2008	2009	2010
Aberdeen, SD	14.878	15.839	18.526	-9.764	-9.451	-11.304
Akron, OH	-9.26	-5.672	-5.795	-5.949	-3.772	-4.351
Biloxi, MS	6.984	5.255	5.227	10.201	0.918	3.809
Boulder, CO	18.698	17.116	21.178	1.335	2.236	1.619
Bradenton, FL	-2.723	2.427	-0.456	5.881	5.251	5.491
Charlotte, NC	-0.790	0.029	-2.872	3.354	2.391	2.717
Columbia, SC	-0.151	-3.679	-2.390	2.983	2.781	1.603
Columbus, GA	1.542	-2.549	0.389	4.888	5.380	9.261
Detroit, MI	-16.113	-13.464	-13.698	-10.425	-8.743	-9.673
Duluth, MN	9.229	5.668	7.561	-5.914	-6.563	-3.368
Fort Wayne, IN	-4.870	-3.629	2.014	-4.245	-0.963	-1.548
Gary, IN	-26.366	-28.762	23.425	-11.648	-11.450	-14.149
Grand Forks, ND	21.510	22.820	23.305	-10.833	-9.007	-6.830
Lexington, KY	4.612	4.540	1.887	1.371	2.817	1.890
Long Beach, CA	1.037	1.097	-1.558	11.918	11.662	10.762
Macon, GA	-8.483	-16.206	-20.061	-0.848	-2.976	-2.441
Miami, FL	-19.206	-16.776	-17.147	3.065	6.186	4.875
Milledgeville, GA	-7.248	-3.406	-15.447	1.288	2.908	0.106
Myrtle Beach, SC	-1.312	-0.552	-5.975	11.443	9.946	8.744
Palm Beach, FL	-13.488	-9.310	-7.918	6.562	7.547	5.194
Philadelphia, PA	-0.255	0.396	-3.360	-2.541	-2.211	-2.399
San Jose, CA	2.188	2.214	0.942	-0.296	-3.231	-5.260
St. Paul, MN	10.302	11.811	11.239	0.837	0.523	2.126
State College, PA	18.993	15.764	26.781	-4.214	-3.728	-1.189
Tallahassee, FL	3.193	4.618	1.787	2.746	2.993	4.395
Wichita, KS	-2.896	-5.588	-0.733	-1.195	-1.403	-0.080

Table 8: Loadings for MDS on z-scores for survey years 2008-2010

Motion charts provide many different ways one can interpret the clusters and dimensions of the Multidimensional Scaling. In each figure, we can see distinct clusters of cities. We can group them by region or urbanicity to search for patterns in the clusters. Dynamic cities, which are those cities that move from cluster to cluster throughout the years, are marked on the charts.

Higher mean scores and proportion scores imply that the city scored higher across all index variables. Dynamic cities with high values in these summary statistics include: Biloxi MS, Palm Beach FL, Milledgeville GA, Boulder CO, Tallahassee FL, Aberdeen SD, Grand Forks ND, Wichita KS, Columbus GA, Long Beach CA, and Myrtle Beach SC. A higher score in standard deviations implies that the responses for that city had more variation across the index variables. Dynamic cities with a high variation in responses include: State College PA, Wichita KS, Tallahassee FL, Palm Beach FL, Philadelphia PA, and Myrtle Beach SC. Higher z-scores indicate a higher city score relative to the original index variables. Columbus is the only dynamic city for the z-scores that scored higher relative to the original index variables.

Multidimensional Scaling has allowed us to identify not only distinct clusters of cities that are more similar in their responses to the index variables, but also detect dynamic cities and observe how the characteristics of the cities change throughout time. The following figures highlight these features.



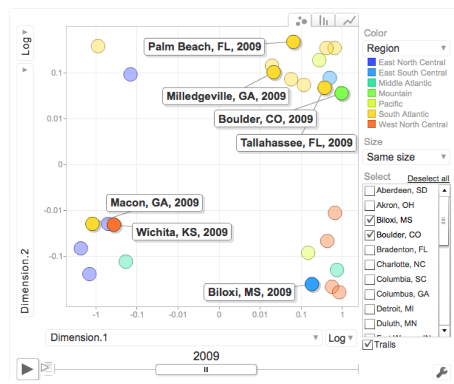
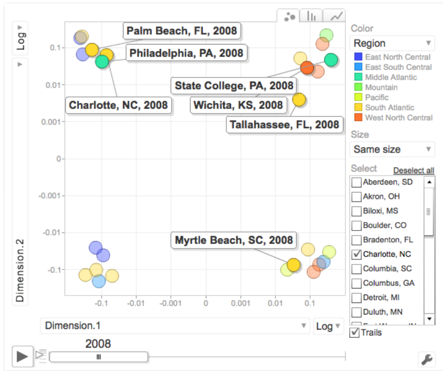
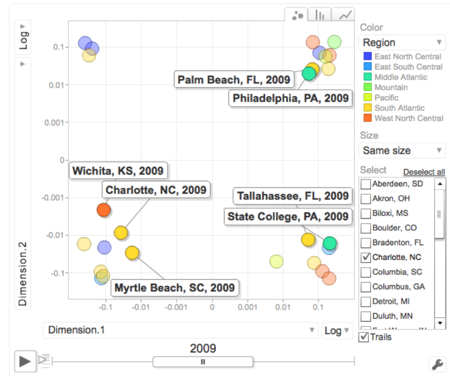


Figure 6: MDS results on the means for 2008-2010. Dimension 1 is on the x-axis and Dimension 2 is on the y-axis. The data are shown on the log-log scale, with color specified by Region





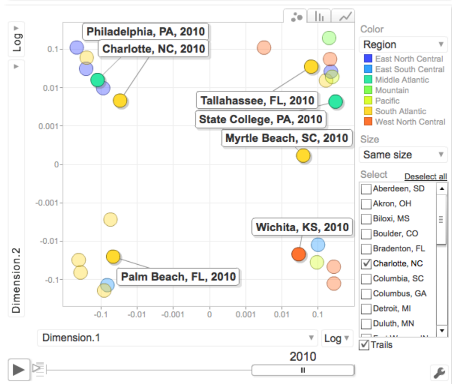
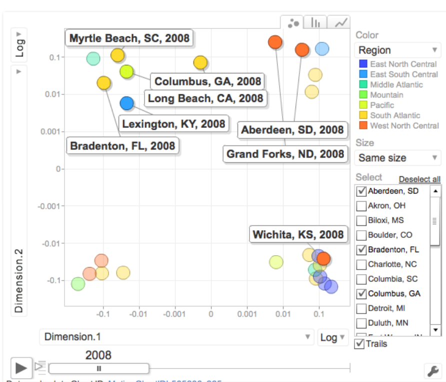
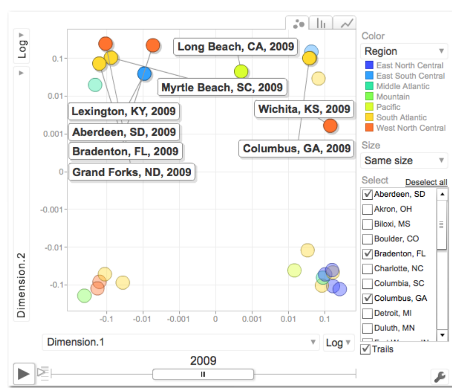


Figure 7: MDS results on the standard deviations for 2008-2010. Dimension 1 is on the x-axis and Dimension 2 is on the y-axis. The data are shown on the log-log scale, with color specified by Region.





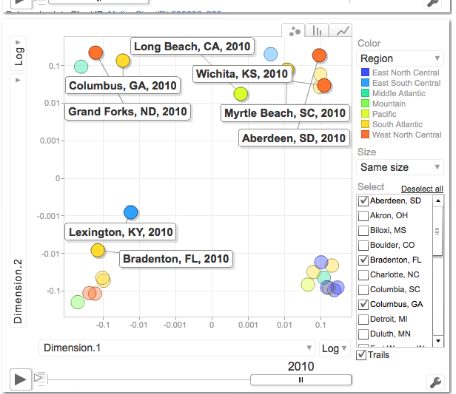
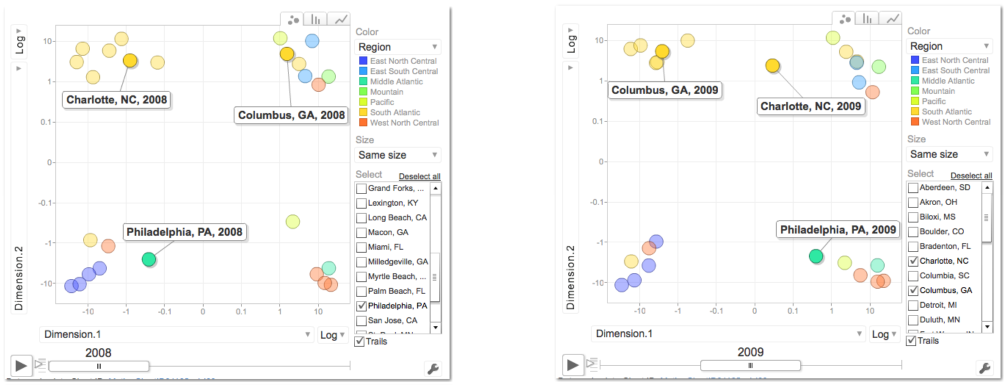


Figure 8: MDS results on the proportions for 2008-2010. Dimension 1 is on the x-axis and Dimension 2 is on the y-axis. The data are shown on the log-log scale, with color specified by Region.



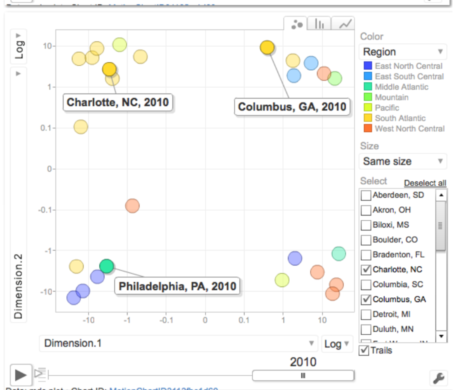
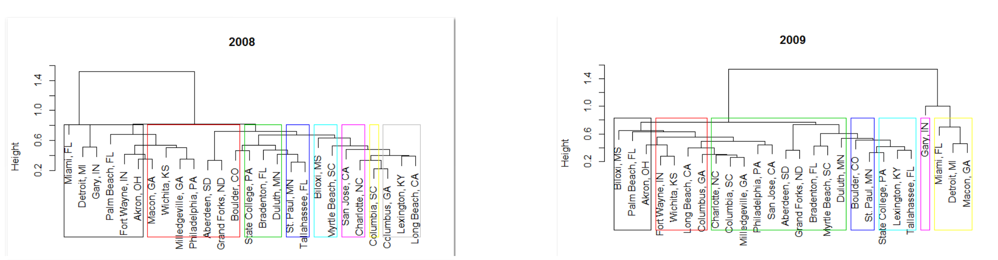


Figure 9: MDS results on the z-scores for 2008-2010. Dimension 1 is on the x-axis and Dimension 2 is on the y-axis. The data are shown on the log-log scale, with color specified by Region.

3.2 Hierarchical Cluster Analysis

Another way we can observe the differences between the communities is to look at the results of average hierarchical cluster analysis. This method seeks to create clusters based on sets of dissimilarities for the cities. Through the use of an iterative algorithm, hierarchichal cluster analysis begins with each city in their own cluster, and then joins the cities together that are the most similar. Figure 10 shows the dendrograms for each year, and the clusters of cities obtained by this method. Cutting each tree at 0.8, we can observe different numbers of clusters for each year, as well as different groupings of the cities throughout time.



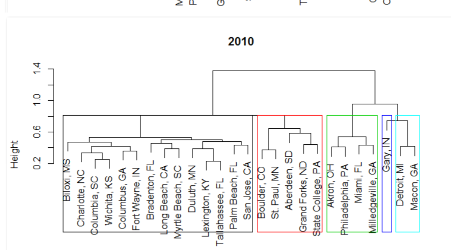


Figure 10: Average Hierarchical Cluster Analysis for survey years 2008-2010

4. Conclusions and Future Research

We have demonstrated the use of motion charts in displaying the results of time-dependent multivariate analysis. Dynamic and interactive interpretations can be achieved and customized based on the interest of the user. Future research in this area will be to repeat the analyses based on subsets of the data by the suggested clusters to further understand the relationships between the index variables and cities, and to better characterize what attaches people to their communities.

REFERENCES

Everitt, B. and Dunn, G. (2001), "Applied Multivariate Data Analysis," *Hodder Arnold UK* .

Gesmann, M. and de Castillo, D. (2011), "Using the Google Visualization API with R," *The R Journal*, 3(2), 40–44.

Husson, F., Josse, J., Le, S. and Mazet, J. (2013), *R package version 1.24. http://CRAN.R-project.org/package=FactoMineR*.

Knight Soul of the Community (2013), *What Attaches People to Their Communities? http://www.soulofthecommunity.org/*.

R Core Team (2013), "R: A language and environment for statistical computing," *R Foundation for Statisticcal Computing, Vienna, Austria http://www.R-project.org/*.

Wickham, H. (2009), "`ggplot2`: elegant graphics for data analysis," *Springer New York* .