

CC: 96-45
RECEIVED

OCT 01 2002

State Cost Cluster Analysis

Introduction

Federal Communications Commission
Office of Secretary

Cluster analysis could be used for identifying urban and rural areas and high- and low-cost areas, or states for purposes of providing support.

Methodology

Cluster analysis is a univariate (or multi-variate) analysis technique that seeks to organize information about variables so that relatively homogeneous groups or "clusters," can be formed. Each cluster thus describes, in terms of the data collected, the class to which its members belong. The classification will depend on the particular method used. It is possible to measure similarity and dissimilarity in a number of ways.

A clustering method is needed to identify clusters of states with high and low costs, so that support can be provided to the high cost states. The univariate Euclidean distance method is used. It measures the distance between two values as the arithmetic difference, i.e., $value1 - value2$.

The method is implemented in the following manner. First, states are sorted by cost per loop (CPL) and the difference is computed. See attached spreadsheet. Next, a series of clusters are computed based on cluster split differences ranging from 2.5 to 0.5. A cluster is indicated if the difference in CPLs between adjacent sorted states is greater than the cluster split difference. Finally, the stability of the clusters is examined over the range of cluster split differences.

Results

The first state to break out into a separate cluster is Mississippi. The second jurisdiction to break out into a separate cluster is District of Columbia. These jurisdictions may be viewed as outliers. The first group of states to break out into a separate rural high-cost cluster are Kentucky, Maine, Alabama, Vermont, Montana, West Virginia and Wyoming. The remaining states ranging from New Jersey to Nebraska form an urban, low-cost cluster. If the two outliers are combined into these two larger clusters, then cluster stability is achieved for a wide range of cluster split differences ranging

No. of Copies rec'd
List ABCDE

0142

from 2.5 to 0.85. Reducing the cluster split differences further (e.g., 0.85 to 0.5) results in fragmentation of states into many unidentifiable clusters.

Interpretation

The U.S. has a large urban cluster that includes the national average, and an 8-jurisdiction, rural cluster that falls outside the national average. The FCC's current method uses these natural groupings as identified by a univariate cluster analysis. The current benchmark of 135% to national average (CPL = \$29.60) provides support to those states in the rural, high-cost cluster. Hence, the FCC's current methodology is providing support to the group of states that can be analytically identified as high-cost or rural.

STATE COST CLUSTER ANALYSIS

STATE	State CPL	CPL Difference	Cluster Split Differences							
			2.50	2.00	1.55	1.50	1.00	0.85	0.50	
Dist. Of Col.	16.0313									
New Jersey	18.09308	2.061777043		*	*	*	*	*	*	
California	18.36228	0.269202909								
Massachusetts	19.18045	0.818170743								*
New York	19.546	0.365544777								
Nevada	19.71822	0.172222482								
Florida	19.86697	0.148755428								
Maryland	19.90223	0.035252677								
Rhode Island	19.99115	0.088927016								
Pennsylvania	20.64201	0.650853312								*
Arizona	20.7334	0.091390437								
Hawaii	20.77071	0.037317385								
Illinois	20.7727	0.001985659								
Utah	21.19423	0.421532696								
Virginia	21.87975	0.685516771								*
Texas	21.90215	0.022404339								
Alaska	22.01712	0.114966706								
Connecticut	22.06188	0.044759922								
Georgia	22.14027	0.078388394								
Minnesota	22.25597	0.11569615								
Washington	22.31244	0.056473756								
Wisconsin	22.71888	0.406439276								
North Carolina	23.18681	0.467928123								
Colorado	23.35272	0.16591541								
Ohio	23.36297	0.010249713								
Oregon	23.41386	0.050887279								
Michigan	23.50088	0.087019363								
North Dakota	23.97012	0.469238239								
Indiana	24.18936	0.219238314								
Iowa	24.28756	0.098201185								
Kansas	24.71672	0.42916505								
Puerto Rico	24.88505	0.168325153								
Missouri	25.07276	0.187712385								
New Hampshire	25.09483	0.022071128								
New Mexico	25.7026	0.607770941								*
South Carolina	26.05983	0.357227222								
Tennessee	26.37502	0.315191549								
Oklahoma	26.38137	0.006345897								
Louisiana	26.41917	0.037807808								
Idaho	26.9214	0.502222235								*
South Dakota	27.77254	0.851149266						*	*	
Arkansas	27.96557	0.193025237								
Nebraska	28.20475	0.239177138								
Kentucky	29.78325	1.518497802			*	*	*	*	*	*
Maine	30.41798	0.384248815								*
Alabama	31.64293	0.228047511								*
Vermont	32.37634	0.733112884								*
Montana	32.72822	0.157882106								*
West Virginia	33.43617	0.4079446								*
Wyoming	33.71669	0.2805202								*
Mississippi	37.78217	4.065475787	*	*	*	*	*	*	*	*
Nat. Average	21.92357									
135% of Avg.	29.59681									

Note: The asterisks (*) indicate which states have a Cost per Loop (CPL) difference that exceeds the cluster split difference.