Journal of the Minnesota Academy of Science

Volume 27 | Number 1 Article 5

1959

Problems of Delimiting Multiple-Component Regions

Leverett P. Hoag
University of Minnesota, Duluth

GEOGRAPHY

PROBLEMS OF DELIMITING MULTIPLE-COMPONENT REGIONS

LEVERETT P. HOAG
University of Minnesota, Duluth

The determination of regions is as basic to the geographer, perhaps, as classification is to any of the natural scientists. It is, in fact, a type of classification. As with other classifications, regions are more than a mere ordering of data or phenomena. They should order and arrange the data to make them as intelligible as possible. As James has stated, "A system of regional divisions is justified if it illuminates the factors or elements of a problem; it is not justified if it obscures these factors or elements." (James: 1954).

Throughout geographic history there has been much discussion of the meaning of the term region, and not without differences of opinion. There seems to be a rather general consensus among geographers now that regions are not naively given, but are intellectual concepts for better understanding the areal distribution of phenomena. There likely is an even greater consensus about the basic characteristics of regions with almost universal acceptance of the idea that regions must possess some degree of internal homogeneity and also must possess external uniqueness or difference from surrounding regions. However, regions are never entirely homogeneous. The investigator must look for those areas of significant similarities and ignore minor differences as well as minor similarities if they do not coincide with the areas of more significant similarities.

Realizing that regions cannot be exactly homogeneous and that differences between adjoining regions seldom are concentrated in a narrow boundary zone, one becomes aware of the difficulty in delineating regional boundaries. However, though regional characteristics are generally best defined in core areas and are somewhat blurred and transitional in peripheral areas, it is customary and sometimes quite essential to draw rather precise boundaries in these transitional areas. It should behoove the investigator, therefore, to place these boundaries in the most logical and enlightening places.

Regions are of many types based on the number of features used in the regional identification and on the relationship existing between these features. The type of region involved in this discussion is somewhat unique. While regions may be single-feature or multiple-feature,

the type of concern here is a single-feature region but of a feature with multiple components. For example, the region might be a land-use region, a type of manufacturing region, a crop-combination region as reported by Weaver (1954), or one of any other feature that has more than one component. The region is identified by the components which are important in the area. Regional differences exist where certain components lose their importance, or other components gain a position of importance. A type of manufacturing region might be called a steel, chemical, and heavy machinery manufacturing area if these items are important and no other items are of importance. Other types of manufacturing are likely present but are not significant.

The major problem encountered in such regionalization is the determination of important components. This has usually been accomplished by classifying each statistical unit by some method. Once units are classified, regional boundaries can be drawn around blocks of contiguous units that fall in the same category. This assumes that a satisfactory method of classifying such data can be developed and that, once developed, boundaries around blocks of contiguous similar units would be geographically valid. Neither of the above assumptions is necessarily true, but to prove that would take more time than is available here. The purpose of this paper is to point out some of the problems and pitfalls that the investigator must be wary of while trying to perfect a process for regionalizing such data.

While importance or significance of phenomena are not always due to magnitude, let us assume for methodological purposes that importance is based on magnitude. This assumption is arrived at because we need a quantifiable basis and magnitude is easily quantifiable. Any other evaluation that can be quantified should work equally well.

When components are being evaluated on the basis of magnitude, should one consider relative or absolute magnitude? In either case the first or largest component must be important. Succeeding ones may or may not be. Let us assume that we have a situation with widely differing distributions, as follows:

A	B	C
95%	35%	30%
5%	25%	20%
	16%	15%
	11%	12%
	8%	8%
	5%	6%
		5%
		4%

(In these, as in subsequent similar tabulations, the magnitudes of the components of some feature are arranged in rank order in columns for the several statistical units.)

If importance is based only on absolute magnitude, one has only to establish a cut-off value and follow the decision precisely. If we assume that an absolute value of 5% or more is important, then in ex-

ample A above there would be a two-component association. The 5%, however, doesn't look very important following the 95%. In example B the 5% looks more important. At least, it should be rather easy for many to call example A a one-component unit, but it would be more difficult to call example B a one-, two-, three-, four-, or five-component unit leaving out the sixth rank of 5%. Also, if a sharp cut-off is established at some value such as this 5%, what would be the justification for excluding the 4% in example C? Is the 4% so much less important than the 5% or the 6%? Perhaps the 5% is equally important in all three examples. If these are cropland acreages, it likely is to the farmer's pocketbook. It is apparent that both absolute and relative magnitudes have certain advantages and certain disadvantages. How to use a combination of these to avoid the disadvantages is not evident.

Another weakness can be found in many methods in areas where numerous components are present and some account for rather small percentages. If the larger percentages are important, then regional differences frequently will be determined by the inclusion or exclusion of rather minor components—components which on occasion may be too minor to account for significant differences. One example is the boundary on Weaver's (1954) 1949 map which separates northern North Dakota from the rest of North Dakota. Wheat, barley, flax, oats, and hay are considered important components on both sides of the boundary. The difference between the two regions is the corn found south of the boundary; this difference is from essentially no corn north of the boundary to about 1 or 2% south of the boundary. This slight variation in an extremely minor crop hardly justifies regional differentiation.

Still another weakness involves the possible use of rank order in classification. In other words, should a corn-oats-hay crop-association region be considered different from a region where the association and rank order is hay-oats-corn? The samples below show an unfortunate result when rank order is not used.

Adams Co., N. Dak.		Beltrami Co., Minn.		Williams Co., N. Dak.	
69%	wheat	50%	hay	77%	wheat
15	hay	18	oats	11	hay
7	corn	14	flax	4	barley
5	flax	5	corn	4	oats
3	oats	3	wheat	2	flax
2	barley	2	barley	2	corn

Adams County, North Dakota, and Beltrami County, Minnesota, were classified by a method used by Weaver (1954) as six-crop counties with the same six crops in the association; therefore, they would be placed in the same region. Obviously they are much different both in rank order and in relative magnitudes of specific crops. At least one regional boundary should exist some place between these two counties. On the other hand this method placed a boundary between Adams and Williams Counties which are really quite similar even

though rank order differs. Williams County was classified as a one-crop county.

Using rank order as a criterion of classification would not always solve the problem though it would help identify the difference between Adams and Beltrami Counties. Below are three counties in North Dakota that would be difficult to classify by any method, but at least it would seem that they are so similar even with differing rank orders that they should be in the same category.

Bottineau County		Ward County		McLean County	
60%	wheat	56%	wheat	57%	wheat
15	flax	16	flax	16	hay
11	hay	12	hay	12	flax
5	barley	7	oats	8	oats
5	oats	4	barley	3	corn
2	corn	2	corn	2	barley

Note the very strong position of wheat, the secondary position of both flax and hay, and the still weaker position of the other three crops.

Another group of samples follows. Here six counties show each of the possible variations in rank order for the three crops, yet in each case there is a remarkably even distribution of the cropland. One could hardly imagine placing any of these in different classes.

Fillmore Co., Minn.		Houston Co., Minn.		Winona Co., Minn.	
35%	corn	33%	corn	30%	oats
30	oats	33	hay	30	hay
26	hay	30	oats	29	corn

St. Croix Co., Wis.		Juneau Co., Wis.		Todd Co., Minn.	
32%	oats	36%	hay	32%	hay
31	corn	33	corn	30	oats
30	hay	28	oats	29	corn

Again we can note both advantages and disadvantages to both sides of the problem. In some cases use of rank order aids the classification, and in other cases it hinders the results.

Still another problem in this type of regionalizing is the possibility of boundaries being placed through quite homogeneous areas where they are not warranted. In southern Minnesota Weaver's method produced seven categories in a block of eighteen counties with Rock, Martin, Kandiyohi, and Lac Qui Parle Counties at the corners. In spite of this great variation in classification, the similarity was quite apparent. In all eighteen counties corn was the first ranking crop and varied between one third and one half of the harvested cropland; oats was second in all counties varying from 20% to 32%; flax varied from 4% to 18%; hay from 6% to 16%; soybeans from less than 1% to 12%; barley from less than 1% to 10%; and wheat from less than 1% to 4%. While this similarity is not astounding, it is, perhaps, sufficient to warrant elimination of some of the seven categories which

resulted. Other areas with even greater similarity and almost as much diversity of classification have been determined. The case of western North Dakota illustrated above by Adams and Williams Counties is another example of a boundary being placed where none exists in actuality.

SUMMARY:

If one expects to develop regions by classifying the individual statistical units, he must face many problems. Among these are: 1. the basis of evaluation, 2. absolute versus relative evaluation, 3. the likelihood of minor components being the decisive factor, and 4. the possibility of the use of rank order as one criterion of classification. These are knotty problems, indeed, but it is only fair to forewarn the potential investigator that they are more likely to be solved than are the problems of geographic validity of the boundaries resulting from such classifications. Though the classification process be perfected, it is still quite likely that boundaries may be placed where no real difference occurs, and may not be placed where significant differences do occur.

In the past, regions frequently have been determined by a high degree of subjectivity. Objectivity in their creation is much desired, but not at the expense of validity.

LITERATURE CITED

PRESTON, JAMES E. and JONES, CLARENCE F. (eds.) 1954. *American Geography: Inventory and Prospect,* Syracuse, Syracuse University Press.

WEAVER, JOHN C. 1954. Crop-Combination Regions in the Middle West. *Geographical Review* 44:175-200.