FROM NON-DATA TO HIGH RESOLUTION MAP INFORMATION

Avi Degani

Department of Geography
Tel-Aviv University
Ramat-Aviv, Tel-Aviv, Israel

ABSTRACT

AVISYS is a versatile analytical mapping system, or rather a research line —concept, working procedures and software— aimed to perform a, variable-resolution, map spatial analysis. The system holds a special potential with regard to high resolution, highly sensitive, analyses, such as the "Urban Micro Analysis." This feature is described in the context of a variety of applications - notably, in planning and decision making. The system utilizes analytical mapping procedures, seeking the identification and isolation of the actual boundaries --within desired class intervals— of areas of a significant subject -matter homogeneity. Hence the term "Homopleth" map is proposed. The system is suited to be used in conjunction with existing spatially orinted data bases —including the use of spatially unmatched forms-- but is, as well, aimed, in some specific tasks, to be a major tool in the initial creation of such data bases. Solutions to the prevailing state of non-data, which is extensively examined, are discussed in this context. AVISYS mapping utilizes a combination of vector and raster techniques (hence "Vecto-Raster"), with an optional color or black and white mode of printing. When color option is used, it is a two to four process color, executed by means of a Line Printer. An option of automated color separation is attached. Examples of applications are brought from the first volume, "Man and Society in the Urban Environment," of a Tel-Aviv thematic atlas series, in preparation by this author.

THE PROBLEM

Map analysis lies in the very heart of an educated spatial analysis. Map spatial analysis is a process involving three, interfaced, major components: a) data, b) an analytical method - used to transform the data into spatial information, and c) the technique - computer aided, or otherwise, which governs the procedure, with great impact on the resulting product.

Contemporary map spatial analysis is both method and technique rich, but is rather data poor. The lack of data, or the unsuitable forms in which data are commonly available for map analysis, reflect on the quality of the analysis. Data problems coupled with related —sometimes consequent— development of less-than-best strategies for map spatial analysis, become a vicious circle. The present paper seeks to pronounce the nature of this problem more specifically, in conjunction with a proposed particular contribution toward solution, as is presented by AVISYS - an Automated Versatile Information System.

THE STATE OF NON-DATA

The Barriers

The state of "non-data," otherwise termed "the data crisis," was treated in detail in the context of geocartographic applications in geographic information systems (GIS) and planning (Degani, 1977, 1978, 1984). The problem, having a general critical effect on the study of space, received some attention in the literature, since sometime (see, for example, Hagerstrand, 1967), yet insufficient attention, seemingly, if to judge by the present state of the art:
A bottle-neck of data problems affects detrimentally the potential of applied map spatial analysis. For the ordinary researcher, engaged in small scale research or detailed planning, must confront, regularly, one or several of the following:
a) There are no data. b) Existing data are not up-to-date or are unreliable. c) Available data are not in an appropriate spatial level of resolution. d) Avialable data, which are intended for comparative study, simultaneous treatment, or analysis of co-variation, exist only in unmatched levels of spatial resolution, or/and different spatial arrangments. e) Available sets of data are, most commonly, choroplethic, while the more efficient —and of higher potential— analytic approach, would lead to adisaggregated map spatial analysis, oriented towards point data, and high resolution results.

The consequences

This, in turn led to an outcome of three major characteristics:
A. In general, the state of the art is that our analytical ability exceeds, by far, our compilation potential. A set of data, consequently, is not a regularly available comodity, which is: a) obtained by the researcher through compilation, b) in the context of a specific problem, c)for use in a particular map spatial analysis, d) aimed at providing applied research specific solutions. But rather, it is often an incidental set of numeric scores, obtained somehow, very commonly for the purpose of demonstrating how a new analysis (algorithm) works (very likely, repeatadly, with the same set).
Furthermore, most literature dealing with data matters, such as methods for data structuring in relation to computerized GIS. But only seldom, primary matters are addressed, such as how to obtain these unavailable data, by direct compilation, or rather, how to create them - in ways of utilizing advanced computerized methods and techniques for map compilation first - prior to map analysis.
B. A great deal, perhaps most, map analysis reported, is of a choropleth nature, or is based on choropleth-structured data. There is little —if any— that is yet to be said about the restrictive character of boundaries of areal subunits, which is inherent in their arbitrary, or non-functional, or meaningless nature. Ironically, to a great extent, automated data bases —presumably modern entities, which are continuously constructed at vast costs— serve, by their often choroplethic nature, as barriers on the way of liberation from this relatively

inferior analytic approach.
C. Existing standard forms of data (e.g. most governmental census forms) —which commonly are both choropleth and of low resolution —permit only rarely, it at all, useful applied research and specific, detailed, planning, which lean on modern spatial analysis.

SPATIAL STRUCTURES AND ANALYTICAL MAPPING APPROACHES

The Conventional, Convenient, Choropleth

As mentioned before, the most commonly found spatial structuring of data, and matching mapping approach, are the choroplethic. Examples of this in big systems would be the DIME (U.S. Department of Commerce, 1970), which is based on variably sized and shaped, city blocks - a classical choropleth; the IMGRID, Harvard University system (Sinton, 1976), which uses regular, equal-sized grid cells, or the San Diego system (San Diego, CPO, 1975), which utilizes a fixed-resolution gridcell data referencing system - to mention just a few out of very many systems.
Alonside such better known systems, which contain both data bases and mapping packages, that are dedicated to these systems and are based on them, there exist numerous mapping programs and packages to handle one form or another of areally-based —choropleth— data structures. The resulting typical map analyses can, obviously, be: (a) never but generalized portrayals, matching the generalized level of the data, (b) never of high resolution, (c) commonly, of variable resolution (between areal subunits) throughout the study area - as is typical of choropleth data structures.

The Dasymetric Option

In principle, a dasymetrically structured analysis and map representation, would be considered an advanced departure from the simple choropleth, and hence a step forward. For dasymetric mapping, as Robinson explains (1963, 75), although based on "exactly the same initial data used for the choropleth map, assumes the existance of areas of relative homogeneity" within the general choropleth pattern, which may be delineated if additional knowledge (data) to support it exists.
The problem, of course, again, is the common lack of such additional data. Consequently, the construction of dasymetric maps —or rather the study of space, employing a dasymetric-type map analysis— remains rather quite theoretical, if to judge by the amount of reported research to that effect. Dasymetric, is thus, the "unmapped" map. So much so, that even cartography text-books show only sparingly examples of dasymetric maps - often, in fact, only schematically worked out small samples.

It is valid to note at this point that although the amount of utilization of the dasymetric map —described by Robinson, and others, so many years ago— leaves so much to desired - due to our continued state of non-data, there have been developed since, modern quantitative mapping approaches, providing significantly more advanced analytical mapping options. Only these are even more sparingly utilized, for their even higher demand for more, and more detailed, data.

The Nature of an Optimal Solution

The inavitable conclusion is that the key to an advanced map spatial analysis is provided by high quality, high resolution, data.
The desired scheme of an advanced, high resolution, map analysis, is one which would —in accordance with the conclusions from the foregoing discussion-- satisfy the following three conditions: a) be totally disengaged from any limiting dependence on any areal subunits (i.e. be non-choropleth); b)present an analytic approach providing for a defendable isolation and delineation of significant boundaries in space of the phenomenna studied; c) have a potential to perform this map analysis in variable scales and levels of resolution - theoretically, any level.
The latter property would serve to analyze a theme in different levels of detail under different circumstances or subject to varying functional assumptions. As well, it would enable analysis of different themes (variables), varying in nature, under different levels of resolution, as may be logically called for.
AVISYS map analysis presents an attempt to implement the above principles, as is demostrated by its governing method of mapping, the "homoplethic."

THE VARIABLE-RESOLUTION "HOMOPLETHIC" MAP-ANALYSIS: THE DASYMETRIC APPROACH REVISED

While dasymetric mapping is basically a choropleth structure; however, one that is elaborated, by additional data, to enabling a refined delineation of the choropleth boundaries; "homoplethic" mapping —a term proposed by this writer-- is a high resolution area-free (non-choropleth) analysis, based on point data.
The governing principle in homopleth mapping is that of the "continuous scanning" (Degani, 1969). Scanning is performed digitally throughout space, in search for point data, which are being screened logically, and are processed geostatistically in a variety of geocartographic methods, leading to the performance of various types of analytical mapping tasks.
The end result is a map (e.g. Fig. 1) depicting the meaningful boundaries —within desired class intervals-- of areas of significant subject-matter homogeneity of the studied phenomenon (hence the name).
In this example (Fig. 1), characteristics of Tel-Aviv's sky-line height are presented as a result of a special type cluster analysis. The map clearly depicts low-structured slum areas and old Arab quarters in the southern parts of the city, and affluent villa areas in the north. As well, shown are the city's uprising building areas in the (new) north, which are purely residential, and non-CBD. The three to four storeyed house is commonly synonymouse in this city, to middle-class residential. Notice the high resolution of the map, which is based on an extensive point sampling. These properties, however, will be subject of a further discussion.

The "continuous scanning" method was f rst described back in 1969, followed by a later detailed description of the basic algorithm (Degani, 1975). Although some seventeen years old, and holding a versatile interesting analytical potential (see, for example, Ron, 1977 - on ancient stone-huts, or Degani, 1975 - on desert streems), the method was used only to a

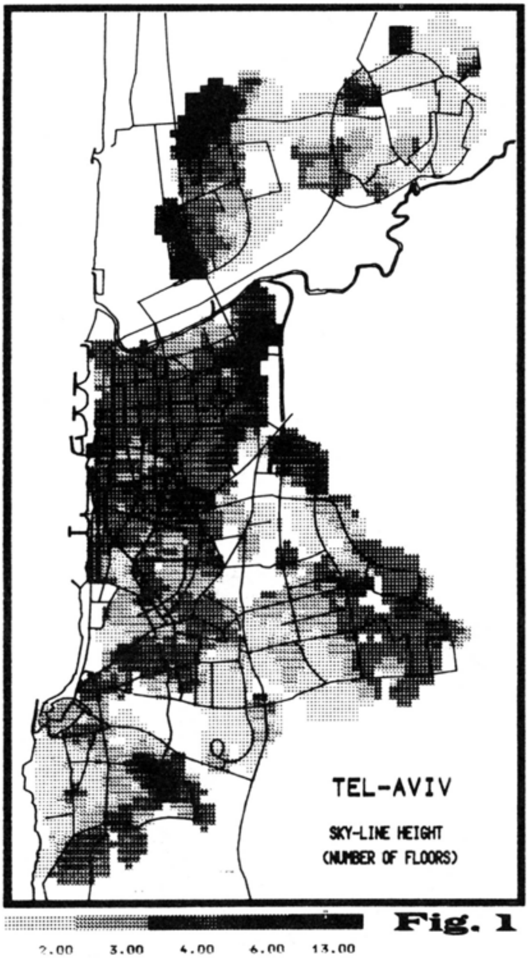
TEL-AVIV
SKY-LINE HEIGHT
(NUMBER OF FLOORS)
2.00
3.00
4.00
6.00
13.00

Fig. 1

relatively limited capacity, mostly because of data limitations. Time again it is evident then, that a high quality map information-content is a derivative of a high resolution map analysis, which, in turn, is critically dependent on the availability of highlt detailed data - practically, point data or close to it.

The inavitably required combination of compilation attached to analysis, has become, consequently, the basic logical structure in the development of AVISYS - software and working procedures.

AVISYS - A VARIABLE-RESOLUTION (MAP) INFORMATION SYSTEM

AVISYS is an analytical system —which is rather a research line— combining analytical mapping software with programs designed for computer assisted map compilation procedures, intended for use in conjunction with field work. The various analytical mapping options of the system (to be specified later) are based on a homoplethic spatial structuring of the map information content. They all utilize detailed point data, and a major feature of their design is the ability to perform a variable resolution, application oriented, map spatial analysis.
To demonstrate: analysis may be as sensitive as to small clusters containing only a few houses in the city - in a "micro analysis" of the urban environment; otherwise, it may generalize greatly - e.g. within areas of several killometers to a side.

The following sections will provide a brief review of these two major components of the system: compilation - carrying the process, initially, from non-data to data, and analysis - that is the transformation of data into various forms of spatial map information.

FROM NON-DATA TO DATA

To generalize, there are three major forms of data of concern in spatial analysis. (a) Physical and social data from census sources, governmental usually. These are typically too generalized - not to speak of problems steming from their common lack of adequate geographic refencing.
(b) Data obtained through remote sensing, in its widest sense. These are very limited, as yet, concerning social data. (c) Social and physical data obtained in the field, both by observation and contact methods (say, polling). These happen to be the most vital sources of data for use in high resolution research and planning. Only until one gets to the field to compile them, one remains in the general state of non-data practically.
In the absence of a break-through to have occured in pollin and interviewing methods or in field observation methods, it is vital to sophisticate the existing methods, in order to bridge the gape, siscussed earlier, between the acquired high state in analytical methods and the lagging low state of compilation methods. This should become a call for the investment of more research energy in the development of computer-assisted compilation (COMACOM ?) to complement —and be part of— computer-assisted cartography (CAC).
This can be done in a variety of ways, all of which have been utilized in both research and practice, related to the development of AVISYS.

In the limited framework of this paper it is possible to touch only very briefly on them:

Exploring Space With a Map & a Computer - Orinted Thought

It is often that we become armchair scholars, although dedicated to the study of space and planning: Because of a fixed state of mind, the compilation in the field of population scores, for example, for an entire city, of, say, 70,000 people, is ordinarily taken to be impractical, or a practical impossibility.
However, if large scale maps are used (e.g. 1:2500) in a scan-by-car procedure of the city, an update on the maps for all existing buildings can be performed (one by one) + count of the number of residential units in each + coding of house-type, all to be completed ... in a single day. This is reported as based on several tested cases of the sort by this writer, for cities in their entirety.
Adding other types of observable items to the count and coding, will increase compilation time, yet not to exceed only a few days, for a couple of people, to compile a rather complex and rich set of data - of the highest spatial resolution there is.
If such maps are then being taken into the lab, digitization of control point coordinates for all buildings may complement the data compiled in the field. This may require a program which can handle the segmentation of the total city map into many large scale sheets. By means of an interactive procedure --which should last no longer than another day or so-- this entire compilation process may end in the creation of an excellent high-resolution data base, already stored onto the computer.
It is hard not to conclude that perhaps the greatest barrier in compilation lies in the human mind.

Computer-Aided, Interactive, Sampling Design

A sampling study may come to complement field, quick survey, methods, as the one described above. Its main purpose would ordinarily be, the obtainment of social and demographic data of all kinds. However, sampling design is more than a matter of pure statistics. When related to spatial analysis, it may involve both geostatistics (much geglected by statisticians) - for a significant design of the spatial allocation of the sample points, and geocartographic modeling - for an actual map allocation of the pattern in the real space.
Such a procedure can result not only in an automatically produced map of the total pattern of the sample population, but can also provide personal maps, enlarged ("zoomed"), for a highly detailed coverage of each field worker's personal section (Degani, 1984).

An Auto-Cartographic Monitoring of Sampling Significance

"Monitorial mapping" while a sampling study is in progress in the field, may be an extremely helpful practice, serving two major goals: a) To examine dynamically the significance and representativeness of the data, obtained through the sampling process, as it comes continuously from the field. This procedure is aimed to minimize sampling size, while testing continuously for possible necessary expansion of it, which is, tentatively, preplanned. b) To search for human errors caused during the field work: Map analysis of variation in the data is performed, using an area-sensitive testing procedure. Areas revealing unexpected variation, or an unexpected magnitude of variation (positive or negative), are checked again, then, in the field.

Probably a great deal more can be done in the development of cartographic computer-assisted-compilation methods. It is advocated to become an important trend in future analytical mapping development, which may, through prompt real-time computer-assisted-compilation procedures, deminish conditions of non-data.

FROM NON-DATA TO INFORMATION

AVISYS employs a variety of analytical mapping methods, all utilizing the homoplethic approach:

A. Densitometry. The measurement and mapping of absolute density (of populations, types of buildings, etc.) is one of the basic options. In itself it provides the user with several variations from which to choose. The output includes rich accompanying statistics, computed from the measured scores. The mapping is performed, optionally, by Line-Printer or by plotter. All mapping is totally flexible in scale. As is now set, the total width limit of Line-Printer mapping (in segments) is 97.5 cm. (38.5") with an unlimited length. These dimensions may, however, be readily lifted and increased to any desired alternative. The plotter mapping is restricted only by the size of the plotter.

B. Comparative Measures. This option contains a variety of matrix comparative measures. For example, comparisons of the resulting matrices of two maps, each being the product of some measure (computed) - all run and mapped simultaneously.

C. Overlay Analysis. This analysis may contain any number of variables, for the entire population, which may be examined, and/or be screened, to isolate a section of the population, which satisfies any number of criteria. Then, the selected population may, further, undergo any AVISYS analysis.

D. Cluster Analysis. This option isolates in space clusters, satisfying a level of "geostatistical significance" (Degani, 1984), according to four criteria: (a) mutual proximity, (b) level (or range) of magnitude (Z values) of the subject matter analyzed, (c) minimum (absolute) number of participating items, which must also present (d) a minimum percentage of the relevant participating population included within the proximity range. Fig. 2 portrays, accordingly, clustering by level of education of head of family. The criteria used were: (a) 200 meters as the level of proximity, (b) levels of magnitude - i.e. total range of years of study investigated: 8 years to 20, (c) minimum number of (sample) points for entry - 2, (d) to represent, at least, 25% of the total population at any given spot, which is to be included in a cluster.

E. Correlation and Residuals Analysis. Commonly, this option is employed by running a series of three maps: the first two execute, each, some kind of AVISYS analysis (usually, the same), the third map is a residuals analysis resulting from an executed map-matrix correlation of the previous two.

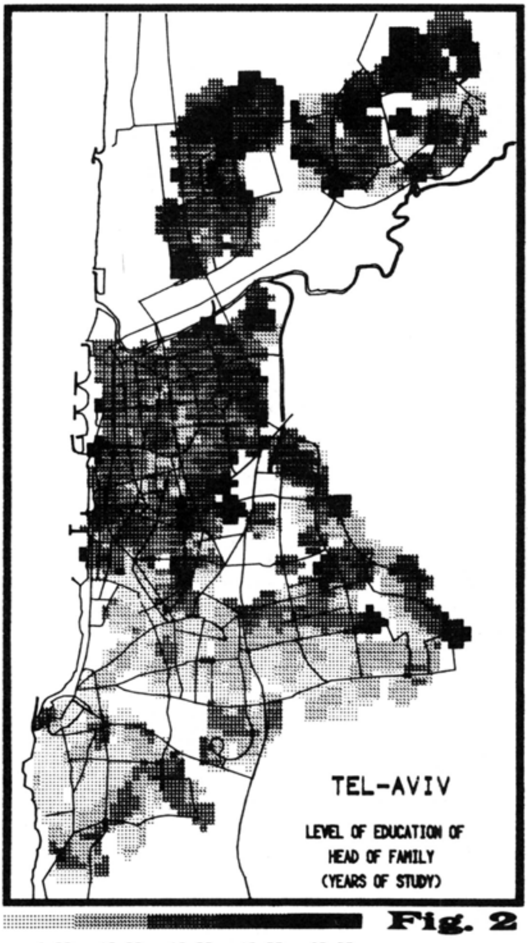
TEL-AVIV
LEVEL OF EDUCATION OF
HEAD OF FAMILY
(YEARS OF STUDY)
8.00
10.00
12.00
15.00
20.00

Fig. 2

A Variable-Resolution Analysis

The ability of the system to perform a variable-resolution homoplethic analysis is one of its more important features. As indicated previously, if permited by the quality of the data, analysis of any of the kinds described above, may be as sensitive as to detect in space homogeneity, or compute area-sensitive geostatistical measures, in reference to small areal pieces, as three or four building clusters. Say, 150-200 meters "immediate neighborhood" level environments. The system could certainly integrate results within any higher level of generalization, or go further down, to a point analysis - which will be, actually, data retrieval. The meaningful level of investigation must, however, be rationalized, in relation to the purpose of the study, by the researcher.

MAP PRODUCTION & REPRODUCTION

Vecto-Raster Color Mapping

All Line-Printer maps are produced, optionally, in color or in black and white. The latter has been imposed in this paper, due to reproduction limitations of this publication.

When produced in color, a process color procedure is used, utilizing an overprinting mode with the Line Printer. Four colors are available: yellow, red, blue and black, as all of their combinations, of course, and ten levels of intensity (of "gray") for each.

The graphics of the map is a combined product of vector and raster techniques, to which the term Vecto-Raster graphics was adopted. In practice, on top of the areal symbolism, first created (raster) by the Line-Printer, all line features and titles are being inserted by the plotter (line) in an interactive, pre-adjusted, matched space, prompt procedure. Thus, the full graphic production of a map of Tel-Aviv, in the scale of 1:50,000 (a logical matrix of some 9000) requires some 7-8 minutes, in three colors.

Automated Color Separation

An option of automated color separation has been added to the color mapping mode (Fig. 3). The great importance of this option, as a budget saver, became particularly appreciable when AVISYS was put to use in the preparation of an extended color publication. A present case in point is the preparation of the first volume of the atlas of Tel-Aviv, which will be discussed later.

APPLICATION

An Urban Micro-Analysis

It is well to stress the particular potential of the system when utilized to perform a highly sensitive analysis, such as in the small-cluster level, described earlier. This high resolution, combined with the potential thematic versatility of map information products, creates high applicability, an important virtue in itself. The detailed, urban micro-analysis may be

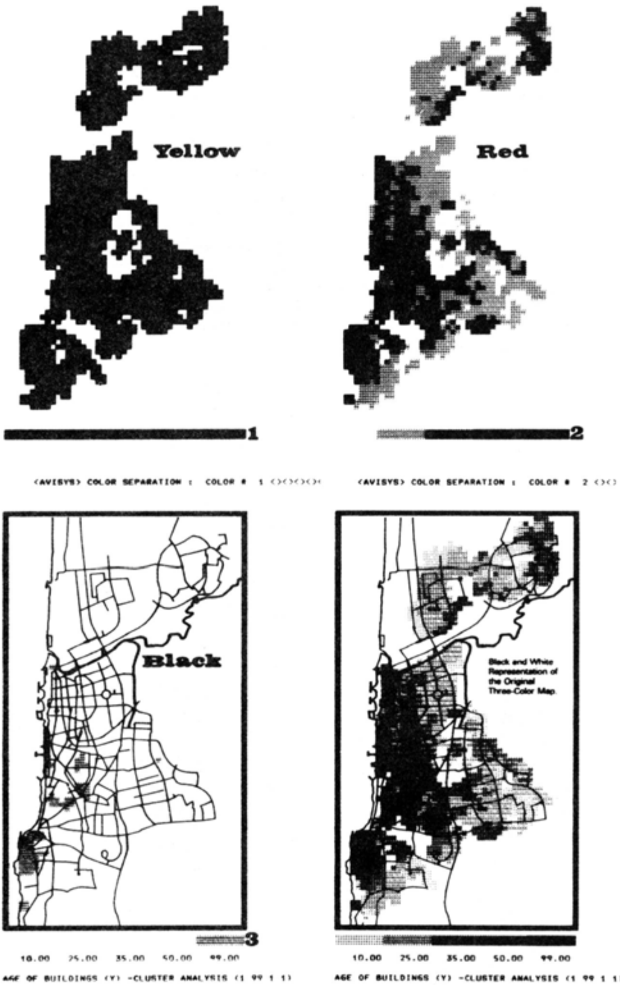
Yellow
Red
1
2
<AVISYS> COLOR SEPARATION : COLOR # 1
<AVISYS> COLOR SEPARATION : COLOR # 2
Black
Black and White Representation of the Original Three-Color Map.
3
10.00 25.00 35.00 50.00 99.00
AGE OF BUILDINGS (Y) -CLUSTER ANALYSIS (1 99 1 1)
10.00 25.00 35.00 50.00 99.00
AGE OF BUILDINGS (Y) -CLUSTER ANALYSIS (1 99 1 1)

Fig. 3

used in reference to the city dynamics, as a prime research tool in conjunction with detailed planning. In routine, it is applicable as a sophisticated supporting means in various circumstances of decision making. As well, a micro analysis may be an extremely revealing means in conjunction with a detailed spatial analysis of results of public opinion polling. The maps nay be used as an educational vivid aid in school teaching of the geographic social and physical structuring of the city complex, and as an extremely rich source of information for use by the general public, as well as by professionals.
For it is very meaningful, for example, in terms of planning a new bus station for the city, to view the very detailed spatial pattern of all daily commuters, clustered by classes relating to the length, in minutes, of their daily travel to work. It is as helpful for officers of the city to receive a sensitive picture of residents' complaints regarding certain central services (e.g. garbage collection), provided by the city. It is, probably, very interesting (perhaps vital) for the mayor of the city, or a mayoral candidate, to get a last-polling city map, showing in great detail the relative strength of his personal charisma versus the strength of his party, or his competitor's, in the coming local election, or all in historical retrospect.

The Atlas of Tel-Aviv

The above examples are all of actual cases, but only a few of some one humdred maps of the sort, which will be included in the Tel-Aviv atlas, a publication in progress. Now in preparation is a first volume, which will focus on demographic and social topics, and will be titled: "Man and Society in the Urban Environment." Due to the, mentioned, highly dynamic, real time capabilities of AVISYS in compilation + analysis + production, the atlas, which is scheduled to appear within the year 1986, will be up -to-date to 1986. It will be based on last year's data, which are used for demonstration in this paper, in addition to fresh data, presently compiled by means of the methods desribed before. This version of the atlas will include a total polled population of 1500 households (some 1.7%). It is believed, that from the technical and scientific points of view, this enterprise may be repeated practically every several months. This is taken to be an encouraging , practical, indication of the professional progress contained in the direction, in compilation + analysis, taken by the development of AVISYS.

REFERENCES

Comprehensive Planning Organization (CPO) of the San Diego Region. (1975) Coordinate and Gridcell Data Referencing System. San Diego, CPO.

Degani, A. (1969) Some Computer and Isodensitracer Applications in Geography. Journal of the Minnesota Academy of Science, 36, 2&3, 104-109.

Degani, A. (1975) ISODEN - Digital Isodensitometry. Computer Applications, 2, 3&4, 375-394.

Degani, A. (1977) Automated Methods of Geocartographic Analysis as a Planner's Tool. ITCC Review, VI, 2, 67-83.

Degani, A. (1978) Spatial Systems and Automated Geocartographic Analysis. Proceedings of the third Jerusalem Conference on Information Technology.

Degani, A. (1984) Advanced Interactive Geocartographic Analysis of the Urban Landscape: Theory & Application. Proceedings of EURO-CARTO III, International Cartographic Association, Graz, Austria.

Hagerstrand, T. (1967) The Computer and the Geographer. Transactions, Institute of British Geographers, 42, 1-19.

Robinson, A.H. (1963) Elements of Cartography. New York, John Wiley & Sons, Inc.

Ron, Z.Y.D. (1977) Stone Huts as an Expression of Terrace Agriculture In the Judean and Samarian Hills. Unpublished Ph.D. Thesis, Department of Geography, Tel-Aviv University.

Sinton, D.F. (1976) I.M.G.R.I.D, An Information Manipulation System for Grid Cell Data Structures. Cambridge, Mass., Graduate School of Design, Harvard University.

U.S. Department of Commerce, Bureau of the Census. (1970) The DIME Geocoding System.