Ilkogretim Online - Elementary Education Online, Year; Vol 20 (Issue 5): pp. 346-357
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.05.36

An Active Cartographic Representation of Colored Choropleth Map in GIS: Populations Density of Old Mosul City Neighborhoods in 2012 as a Model

Omar Abdullah Ismaeel, Geography Department, College of Education for Humanities, University of Mosul, Iraq, omar.a.ismaeel@uomosul.edu.iq

Abstract- The current research aims to apply geographic information systems GIS in investigating cartographic treatments required for colored choropleth maps in order to reach an effective cartographic representation. The research is based on anassumption that GIS can be used to derive an optimal number of categories, color patterns, and statistical classification methods for the mapped categories. The research relied on the experimental scientific method. The research has been divided into four sections: The first sectioncovers the theoretical framework for the research, the second covers the techniques used in the cartographic treatments of colored choropleth maps in GIS, the third sectionpresents the results and discussion, while the fourth and final sectionstatesa number of conclusions, all of which revolve around that GIS users must study the basic rules of cartography, in order to acquire high level of objectivity in their skills in map design and production.

Keywords: GIS, cartographic treatments, Colored Choropleth Map

I. Introduction

One of the most popular types of thematic maps is the colored map displayed cartographically in choropleth. A choropleth map focuses on a single themein which data is related to units of area termed polygons, anda range of category gradientsareusedin data classification.In such maps,categories are symbolized by colors that reflect the quantitative impression of the represented phenomenon. However, representation of choropleth mapsnecessitates several requirements some of which are cartographically artistic, and some are statistical in their nature. Thus, some cartographers take these requirements into consideration, but others usually neglect them, especially when the relay on GIS in creating maps tosimplifytheir design and output.

The research problem stems from the fact that cartographically illiterate GIS users may be unaware of the consequences of representing and producing colored choropleth maps without correct cartographic treatment. Such overlook from their part may lead toerror laden maps, wrong information delivery, wrong interpretation, andpoor cartographical communication. It is expected that GIS users,unskilled in cartography, perform data representation and classification on colored choropleth map using categories that do not follow cartographic rules. Users may fail tochoose the appropriate number of categories and coloring methods that express quantitative data, and statistics that would help them determine the appropriate categorization method.

Stemming from the aforementioned problem, the research was based on the following hypotheses:

1- It is possible to reach an appropriate number of categories thatcorrectly reflect the spatial variation of the represented quantities in colored choropleth maps.

2- For any choropleth map, a suitable coloring technique can be found and which stimulates the optimal visual impression of the represented quantities.

3- For any choropleth map, it is possible to find an optimal way to classify the categories that simulate the nature of the spatial variance of the represented quantities.

The research aims to investigate the necessary cartographic treatments for the coloredchoropleth maps in GIS in order to achieve an effective cartographic representation of these types of maps.

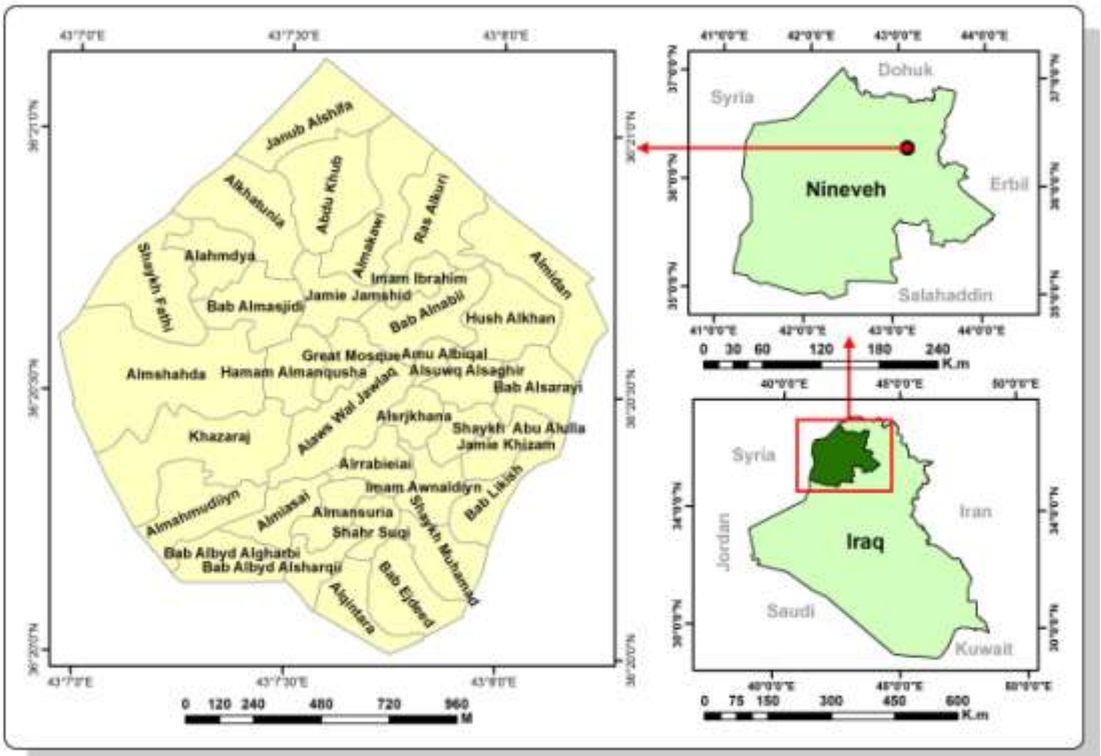
The research relied on the experimental scientific method based oninductive reasoning by making generalizations from specific observations in order to reach an effective cartographic representation of the colored choropleth map.

The Place and time limits of the research were set as follows :

• Place: Thirty-six residential neighborhoods of the old city of Mosul.

• Time: Population density data of the residential neighborhoods of the old city of Mosulin 2012, (see map 1 and appendix 1).

Map (1) . Residential Neighborhoods of the Old City of Mosul in 2012



Source: Iraqi Ministry of Municipalities/ Mosul Municipality Directorate / Geographic Information Systems Division, 2012.

II. Review

The map is asymbolic model of the real world, and map design is the process ofabstracting and modeling the real world throughsymbols().Maps were designed with a fundamental objective reason which is communicating environmental information().

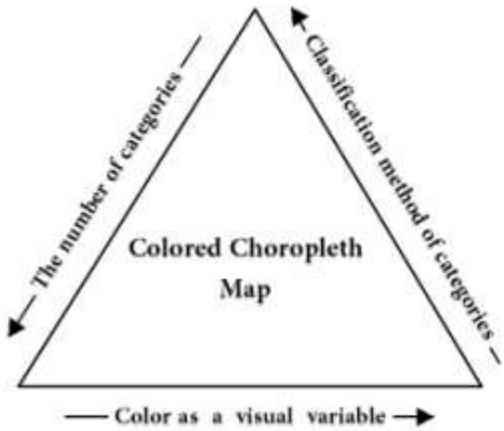
Acolored choropleth map is a miniature visual model of the spatial organization of relative statistical data, and it is one of the quantitative thematic maps that highlight the spatial values spread over a geographic area using colors as a visual variable. In addition to representing the statistical values in each area unit, special quantitative values must be assigned a color gradient pattern in these maps. There are several synonymous to the term ofchoropleth maps.Some cartographers call themrelative distribution maps,density maps, orshading maps. Whatever the name, suchmaps shoulddisplaythe statistical values of a phenomenon in a geographical framework related to the place and not in an abstract way.

In recent years, the use of colors in production of maps including choropleth maps became more common due to the technological developments in both hardware and software. It is beyond dispute that colors have a selective ability that helps to convey the information to the map reader, as it attracts his attention and stimulatesautomatic interaction with the map reader who is used to observe colors.

2.1Considerations in Processing Colored Choropleth Maps

There are three interrelated approaches to process colored choropleth maps. These approaches interact with each other to reflect the spatial variation and the distribution pattern of the geographical phenomenon represented by the map. Therefore, those who seek to design a colored choropleth map in GIS must take the following processes into consideration,(see Figure 1):

Figure 1: The Structure of the Colored Choropleth Map in GIS



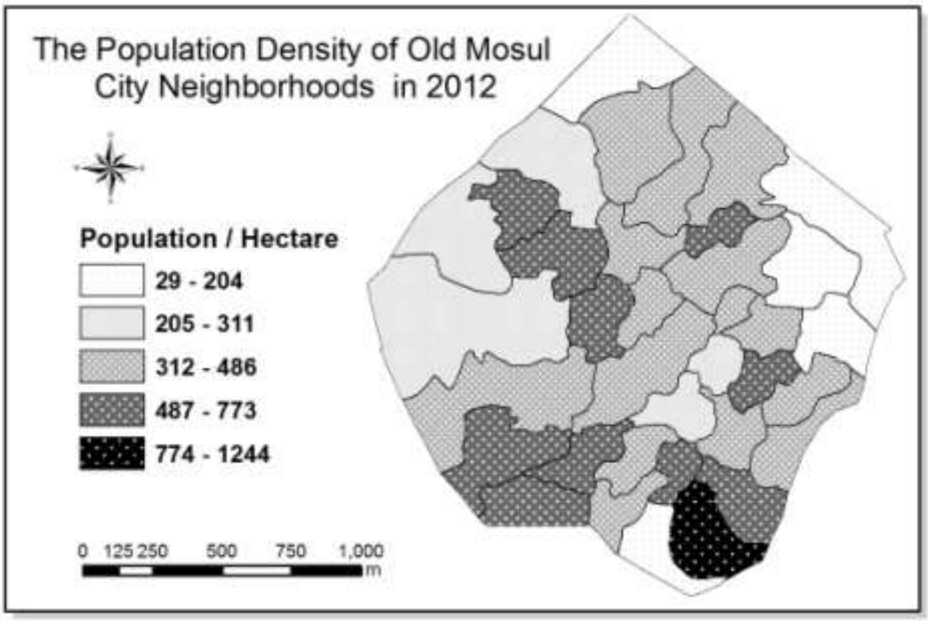
2.1.1 The Number of Categories

Classes in the choropleth map are grouped based on the number of categories. However, the larger the number of categories, the more classes are displayed, and more statistical spatial differences appearin the mapand this makes it difficult to differentiate shadingsusing colors.Thus, the map fails to produce its intended output. The opposite is true whenthe number of categories are low leading to overgeneralization and loss of many details that reflect the nature of the spatial differences inherited in the data.

For this reason, the number of categories used in the choropleth maps has become somewhat standardbecause the map reader cannot easily distinguish between more than (11) area symbols. Thus, it is not recommended to use more than (6)categories and less than (4) categories in producing such maps.The reason behind this simplification is to manage color shading andachieve a more comprehensiblemap ().

Therefore, the colored choropleth maps of the population density of old Mosul city neighborhoods in 2012 were designed in this research using (5) categories as an average between the minimum and highest required number of categories,(see map 2. The design uses the variable of (grain) normalized to the variable of (value) to select and arrange and embody the spatial variance of data on the map (), instead of the color variable while using adefault classification method for the category rate, because the color variable and classification method will be subject to subsequent cartographic treatments.

Map (2): The Appropriate Number of Categories of the Choropleth Map Adopted in the Research.



2.1.2Color as aVisual Variable

Color is an important selective visual variable, readilynoticed by the naked eye and easy to memorize. Humans see colors resulting from the dispersion of white light within the boundaries of the visible spectrum. The process of perception of color in humans produces several representation difficulties to map makers resulting from three color properties termedhue, value and saturation (). Colors are used in mapsas one of the cartographic languages through which a high-efficiency cartographic communication can be established.

The cartographic literature emphasizes the use of warm colors, such as red, and yellow, in the choropleth maps representing population variables such as population density. The origin of the use of color temperature is related to the color legends in heat maps ().

2.1.2.1 Hue

Hue is one of the main propertiesof a color and characteristic by which it is known and distinguished from other colors. Hue is the identity or the name of the color which can be red, yellow, green…etc. The hue of a color can be changed by mixing colors and the pigmentation in the color depends on the color's location on the spectrum or its wave length. Thus,hueis the name used to differentiate colors that have different wavelengths (). The digital value of the huein GIS is determined by a number between(0 and 256).

2.1.2.2 Value

Value, as a color property, refers to the lightness or darkness of a color. It indicates the quantity of light reflected from a specific color . For example, we perceive parts of red colored objectdifferently when half of it is in the shade and the other half in the light.Although the hue of the color is unchanged, we notice a difference in brightness between the two halves. Color value is a mean to distinguishes between the illumination and the darkness of colors, which means that this color is light or dark based on the amount of light reflected by the color ().

2.1.2.3 Saturation

Saturation is the third color property .It is also called (intensity) and it is attributed tothe color's wavelength. The morethe wavelengths of light rays falling on the eye are homogeneous, the moresaturated the color will be and the more the rays differ in their wavelengths the less saturated the color will be. Pale yellow that is hard to differentiate from white is unsaturated yellow , while pink is considered unsaturated red.

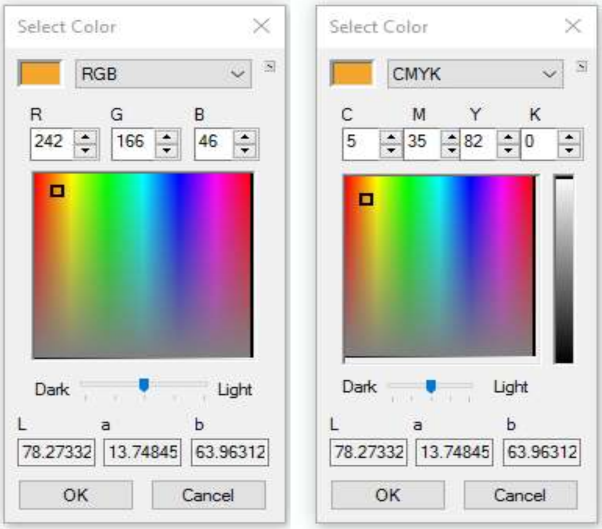
Therefore in color properties, the intensity differs from value.Value is attributed of the amount of light that reflects the color, while intensity is the purity of the color. Therefore,when a color is fully saturated it is considered in its purest form(). The digital value of color saturation is determined in GIS as a percentage and the higher the percentage the higher the color purity and vice versa.

2.1.3 RGB and CYMK Color Modes in Arc GIS Desktop 10.x

ArcGIS Desktop 10.x used in this research relies on the RGB color mode to display images and maps on a computer screen as does other software that processes images and graphics. The RGB mode is an acronym for Red, Green, and Blue,which are the primary colors.Thecolor mode adopted by the software is an optical mode meaning that transmitted light is used to show colors and is only suitable for display screens including those used in computer displays. Therefore, those who choose colors in the maps produced by GIS software use this color mode. For example , the value of the red color in this system is determined by a number that starts from(0) to (256), thedigital number of redmeans greater amount of red in the color and vice versa, as is the case with green and yellow colors.

The CYMK color modeis used for printing mapsafter they are produced using the RGB mode in GIS. The CYMK color mode is the opposite of the RGB optical mode, as it uses reflective light to show colors, so it is adopted in the process of printing maps. The CYMK mode is an abbreviation for Cyan, Yellow, Magenta, and Black, which are colors derived from primary colors. The digital value of the four colors of this mode is determined as percentages in the GIS software, that is, whenever the percentages of the four colors changes in respect to the other colors, it leads to the emergence of a new color (),(see Figure 2).

Figure 2: RGB and CYMK Color Modes in ArcGIS Desktop 10. X



2.2Category Classification Method

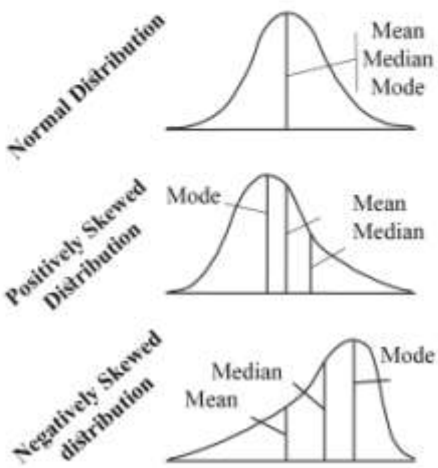
Classification of categories inchoropleth maps is a necessary process aimed at finding classes of data in the form of intervals because the choice of intervals will mainly affect the content of the information represented in the map. Thus choosing the appropriate intervals is a necessary skill for any GIS user (). There are several methods for classifying groups on choropleth maps in GIS. These methods depend on the frequency distribution of statistical data to be represented on the choropleth map.

2.2.1 Data exploration

Data exploration means checking the frequency distribution of data using a histogram. The histogram helps to depict the statistical distribution of a given variable, while the map helps to understand its geographical distribution.Both tools are interrelated. Thus,choosing the appropriate method for classifying the categories depends on the histogram that shows the shape of the values' distribution().

Data frequency distributions take three forms; normal distribution,positivelyskewed distribution, and negatively skewed distribution. The data in normal distribution appear as a curve which peaks in the center while the tails decreases symmetrically, and thus the mean,median and mode are equal ().A positively skewed distribution indicates the data are more frequent at the lower values and the mean is greater than the median and mode. This is reflected in the frequency curve by the peak of the curve appearing on the left. A negatively skewed distribution indicates that the data are more frequent at higher values which means that the mean is lower than the median and mode. This is reflected on the frequency curve by the peak of the curve appearing on the right (), (see Figure 3)

Figure 3: Frequency Distributions of Data



Accordingly, there are several methods of categorizing data in choropleth maps as follows :

1. The Equal Intervals Method converts the values into equal interval categories regardless of the number of spatial units.Equal intervals are calculated by dividing the difference between the highest and

lowest values by the number of categories (), and this method is only valid with the normally distributed data.

2. The Quintile Intervals Method creates categories that contain equal numbers of features.If (5) categories are selected in this method to classify data consisting of (100)features, (20)features are assigned to each category (), and thus this method is only suitable with moderately distributed data.

3. The Geometric Intervals Method is used when the nature of the data distribution takes the form of geometric series. The logarithm of attribute values is calculated for each of the features, and category intervals are determined on the basis of the extracted logarithm values ().

4. The Natural Breaks Method was developed by the cartographer George Jenks and is based on categorizing according to the clusters and gaps in the data(). In other words , it is based on the natural grouping inherent in the data.The natural breaks method reduces the variancein the attribute values in each category and increase the variance between the categories in general(), in other words, the natural breaks method achieves the highest homogeneity in the values within one category, and the highest variation in the values between the different categories, and can also be used with all frequency distributions().

III. Methodology

3.1 Testing the Visual Perception of Color inChoroplethMaps

In order to test the readers visual perception of color and arrive at the best cartographic representation of the color variable in the coloredchoropleth map, ArcGIS Desktop 10.x was used in designing four model maps based on the change in the three color characteristics, namely: value, hue, and saturation in the value of the category data as follows:

1. Hue is variedwhile value and saturation are kept constant.
2. Saturation is varied while hue and value are kept constant.
3. Value is varied while hue and saturation are kept constant.
4. All color characteristics are varied,(see Table 1).

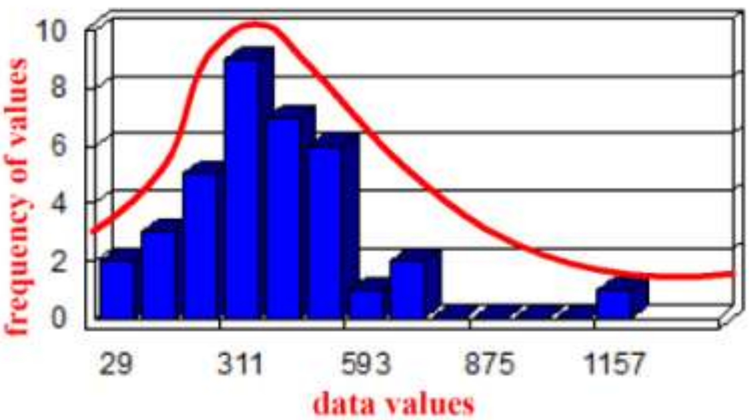
Table (1): Color Modes to be Tested in a Choropleth Map in GIS

Hue is varied while value and saturation are kept constant										
Category	Characteristics of Color-Category			RGB Color mode			CYMK Color mode			
	V	S	H	B	G	R	K	M	Y	C
29 - 272	%100	%100	60	0	250	250	%0	%0	%90	%0
273 - 515	%100	%100	80	0	250	175	%0	%90	%0	%23
516 - 714	%100	%100	216	250	108	0	0%	%65	0%	%90
715 - 1001	%100	%100	0	0	0	250	%0	90%	90%	5%
1002 - 1244	%100	%100	297	250	0	242	%0	%90	5%	%5
Saturation is varied while hue and value are kept constant										
Category	Characteristics of Color-Category			RGB Color mode			CYMK Color mode			
	V	S	H	B	G	R	K	M	Y	C
29 - 272	%100	%20	60	208	250	250	%0	%0	10%	%0
273 - 515	%100	%40	60	159	250	250	%0	%0	25%	%0
516 - 714	%100	%60	60	108	250	250	%0	%0	45%	%0
715 - 1001	%100	%80	60	55	250	250	%0	%0	85%	%0
1002 - 1244	%100	%100	60	5	250	250	%0	%0	%100	%0
Value is varied while hue and saturation are kept constant										
Category	Characteristics of Color-Category			RGB Color mode			CYMK Color mode			
	V	S	H	B	G	R	K	M	Y	C
29 - 272	%100	%50	60	128	250	250	%0	%0	%53	%0
273 - 515	%80	%50	60	102	210	210	%0	%20	%65	%25
516 - 714	%60	%50	60	77	155	155	%0	%40	%74	%43
715 - 1001	%40	%50	60	51	107	107	%0	%60	%81	%64
1002 - 1244	%20	%50	60	26	55	55	%0	%80	%92	%81
All color characteristics are varied										

Category	Characteristics of Color-Category			RGB Color mode			CYMK Color mode			
	V	S	H	B	G	R	K	M	Y	C
29 - 272	%100	%50	60	128	255	255	%0	%0	%50	%0
273 - 515	%98	%66	45	85	209	250	%0	%18	%67	%2
516 - 714	%95	%81	37	46	167	242	%0	%35	%82	%5
715 - 1001	%68	%89	25	19	83	173	%0	%67	%39	%32
1002 - 1244	%42	%100	0	0	0	107	%0	%100	%100	%58

The test was carried out on a sample of(15) graduate students in the departments of geography who have knowledge of reading maps, by showing them four maps while hiding the map legend to ensurethat the cartographic representation of the colored choropleth map is be realized in a visual moment.
Referring to the map legend and reading the change in population density while comparingthe visual impression of each colored group with the colors of the spatial units in order to make a decision for correct perception will take a long time to read the map. This contradicts with the goal of effective cartographic representation of maps which is to ensure immediate perception of the map at the moment ofvisualization. For this reason the map legend was hidden from the participants. The participants were required to answer five questions based on five categories hidden in the maps provided to them, so that each correct or wrong answer received a value of (20%) of the total score of (100%).
The questionwere as follows : -
1. What is the color that belongs to the first category which represents the lowest population density?
2. What is the color that belongs to the second category, which represents the population density higher than the first category?
3. What is the color that belongs to the third category, which represents the population density higher than the second category?
4. What is the color that belongs to the fourth category, which represents the population density higher than the third category?
5. What is the color that belongs to the fifth category, which represents
the highest population density?
From a statistic point of view , the population density data of old Mosul neighborhoods in 2012 werestatistically explored and appeared to have positivelyskewed distribution ,(See Figure 4).

Figure 4: Frequency Distribution of Population Density Data of Old Mosul Neighborhoods in 2012 using GIS



IV. RESULTS AND DISCUSSION

4.1 Results of the visual perception test for the color variable on the Choropleth map.
The maps used in the aforementioned test were labeled(A, B, C, and D) and the results were as follows:
1. In map (A) in whichhuewas varied while value and saturation were kept constant,only (7)participants gave the correctof answers for the first, second, and third category,representing(46.5%) of the total sample who gave(60%) of the correct answers. These participants depended on their experience in interpretation of the characteristic of color and interpreted yellow as shallowness in the first category, green as abundance in the second category, and blue asrichness in the third category. As for the fourth and fifth categories, representedby red and violet, the answers received from all the participant were wrong,

meaning that(100%)failed in the interpretation of the color depicting these categories. In other words, almost half of the sample read the map with an error rate of (40%), and the other half read the map with an error rate of (100%). By adding the two error rates for all the participants in the sample and dividing them by (2), we determine that the average error rate in the visual perception of the total tested sample is (70%)when using hue for cartographic representation of the colored choropleth map.

2. In map (B) in whichsaturation was varied while hue and value were kept constant,all participants gave the correct answers for the first, second, and third category, representing (100%) of the total sample who obtained a degree of (60%) of the correct answers. As for the fourth and fifth categories represented by red and violet, the answers received from all the participant were wrong, meaning that (100%) failed in the interpretation of the saturation. This result indicates a clear confusion in distinguishing between the difference in the amount of saturation. In other words, the average error rate in the visual perception of the total sample tested when using saturation as the basis for the cartographic representation of the colored choropleth map is (40%).

3. In map (C) in whichcolor value was varied while hue and saturation are kept constant, the number of correct answers for the first, second, third, fourth and fifth categories was recorded by (13)participants. These participant represent(86.5%) of the total sample who were successful in perceiving the spatial variation of population density in the map in one visual moment witha total correct answers score of (100%)and error of (0%), while two participants representing(13.5%) of the total sample gave wrong answers in perceiving the fifth category depicting the highest population density. These two participants scored (80%) of the total correct answers with an error rate of (20%). In other words, the average error rate in the visual perception of the total sample when using the color value as a basis of cartographic representation of the colored choropleth map is (10%).

4. In map (D),the three color characteristics; hue, saturation, and value were varied. All participant representing (100%) of the total sample gave correct answers for all categories with a (0%) error rate for map (D). The participants were able to answer all the questions easily and clearly, (see Figure 5 and 6).

Figure(5):Applying Color Modes on GIS Experimental Choropleth maps

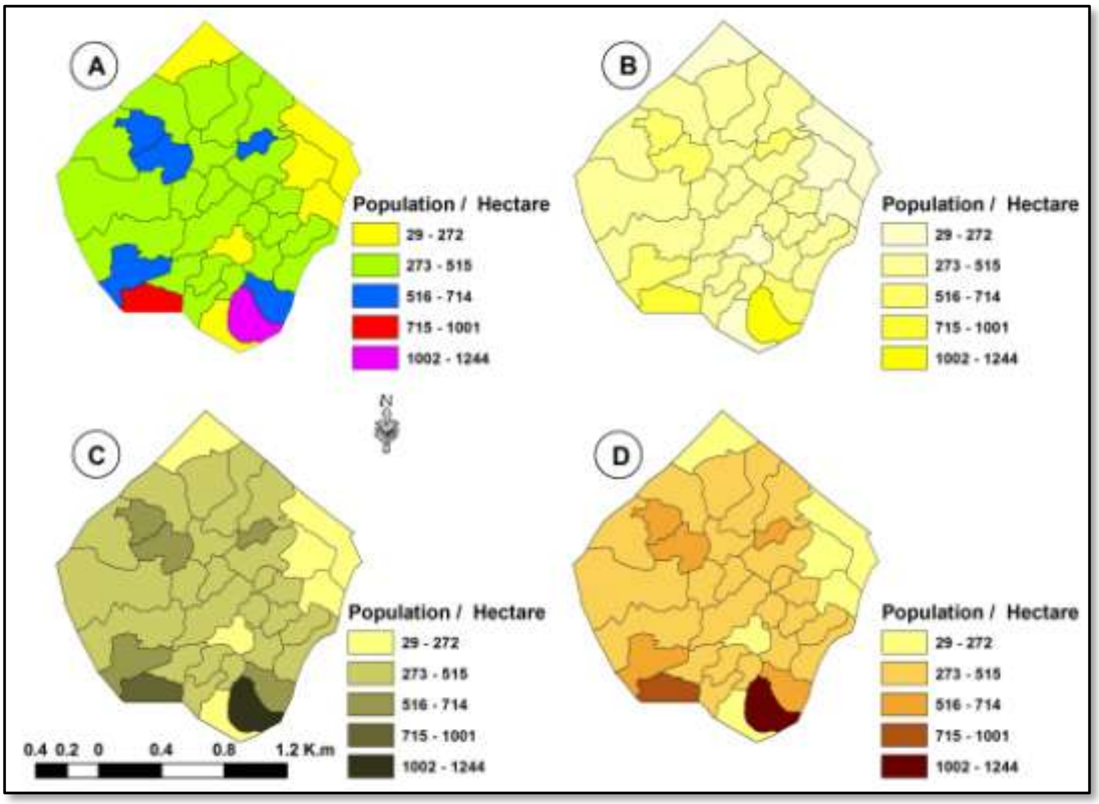
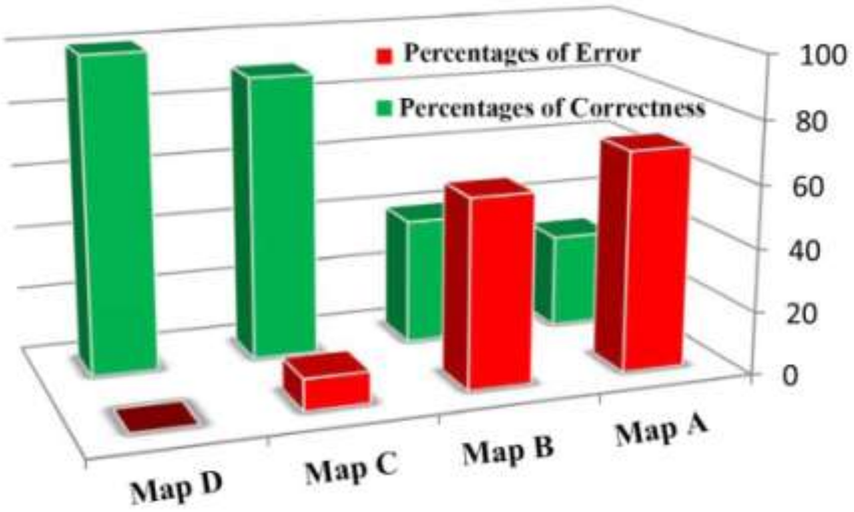


Figure (6): Percentages of Correctness and Error of Visual Perception of the Color Variable on the GIS Experimental Choropleth Maps.



Map (D) achieved a visual perception of (100%). When the three color characteristics were all varied, they effectively reflect the spatial variation in population density between the spatial units in the choropleth map, thus achieving the goal for which the map was prepared as a means of effective communication between its producer and readers.

4.2 Results of the Category Classification Method Test on the ChoroplethMap

Four maps were designed based on the statistical categories'classificationmethods and these maps were (labeled A, B, C and D), respectively as shown in Figure (7).

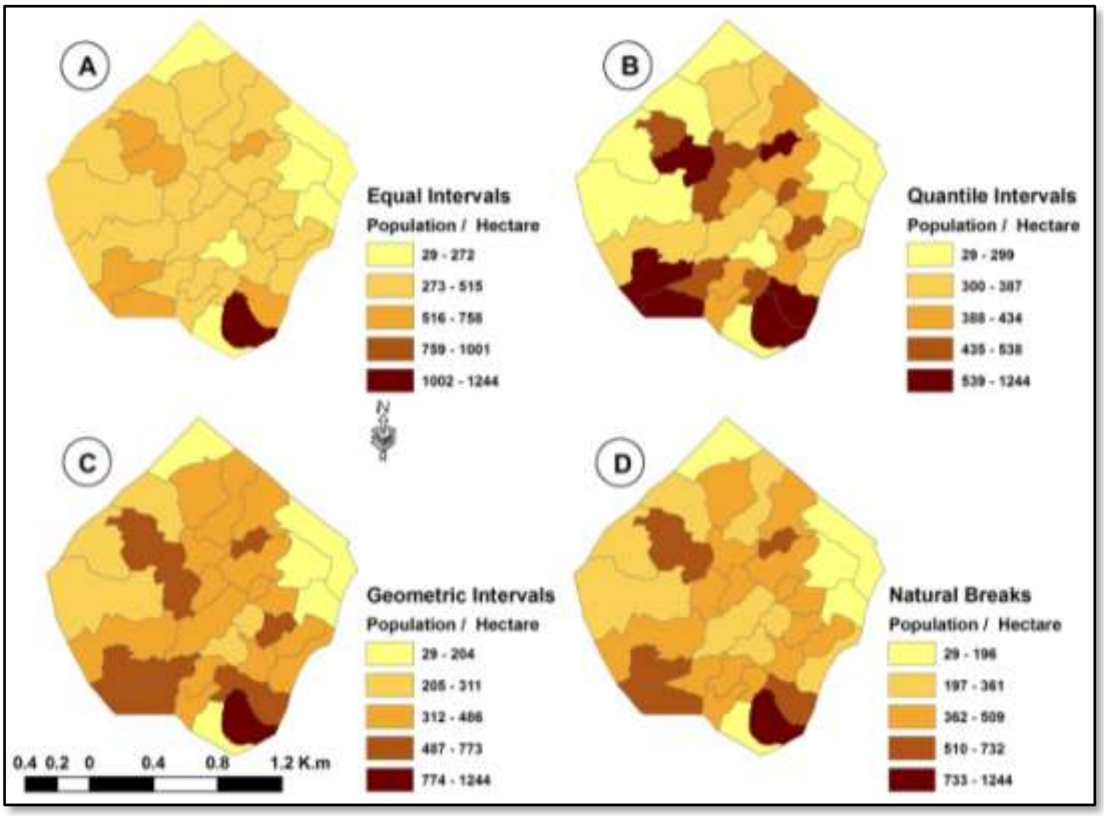
1- Map (A): The data on the map represented by the equal intervals method was a misrepresentationof reality.Population density in this method appeared as if they were close invalues. The reason behind this misrepresentation is because the equal interval method works with the natural distributed data, while the data adopted in this research arepositivelyskewed in their distribution .

2- In Map (B): the population density data were represented usingthe quantile interval method and the map shows large variance between residential neighborhoods. The reason behind this misrepresentation is because this method works to divide the data into arbitrary periods regardless of the inherent variation in the data.

3- In Map (C): The population density data were represented using the geometrical interval method and appeared somewhat close to reality because this method is appropriate with skewed data distribution method, but was imperfect due to the nature of this method that works to extract the logarithm of the data to get rid of anomalous values. Since the data adopted in this research do not contain anomalous values, this method led underrate the spatial variation of population density in the map.

4- In map (D): The population density data were represented using the natural break method and appeared very close to reality, because this method works for all types of frequency distribution of data except for data that carries large differences in values.

Figure7: Applying Categorizing Methods to GIS Experimental Choropleth Maps



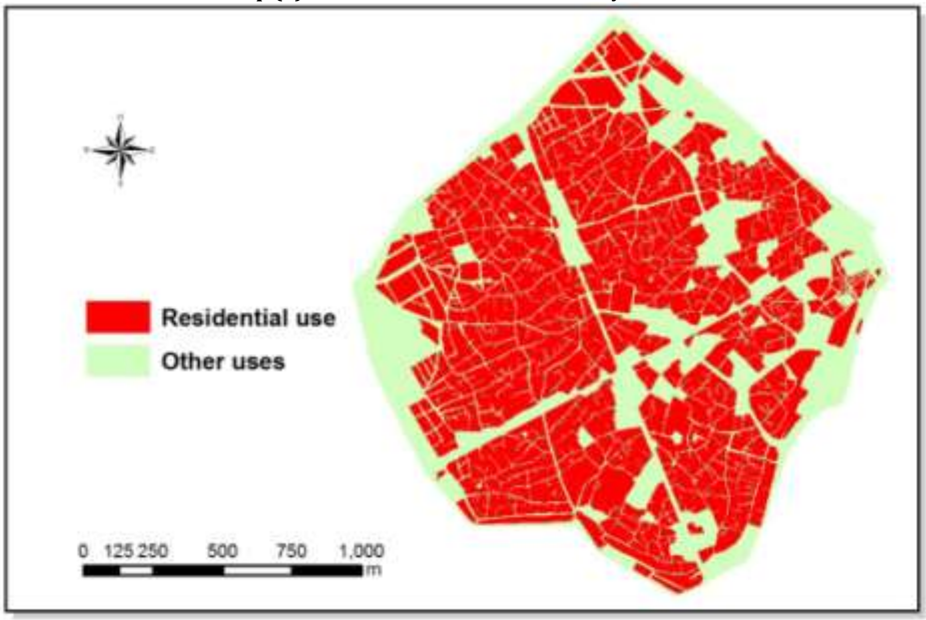
V. THE MISREPRESENTATION PROBLEM

A population density choropleth map is a cartographic representation when mapping the population per area in each spatial unit regardless ifthe spatial unit has uses other than residential use.These non-residential areasare included in population density calculation although they are unoccupied leading to overgeneralization that in turn produces misleading information about the spatial variation of the population density distribution shown in the choropleth map.

In order to overcome the generalization problem in the population density choropleth map of the old city of Mosul in 2012, the following measures have been taken:

1. The population density choropleth map of the old city of Mosul in 2012 was overlaid with the residential use of the same area.
2. Only the areas of occupied residential neighborhoods that fall within residential use were recalculated.
3. Population density is recalculated by dividing the population by area of the occupied residential areas based on the residential neighborhoods to obtain the actual population density for each residential area.
4. The new extracted population density values were reallocated to the occupied residential neighborhoods to obtain an effective cartographic representation of a colored choropleth map free of generalization,(see map 3).

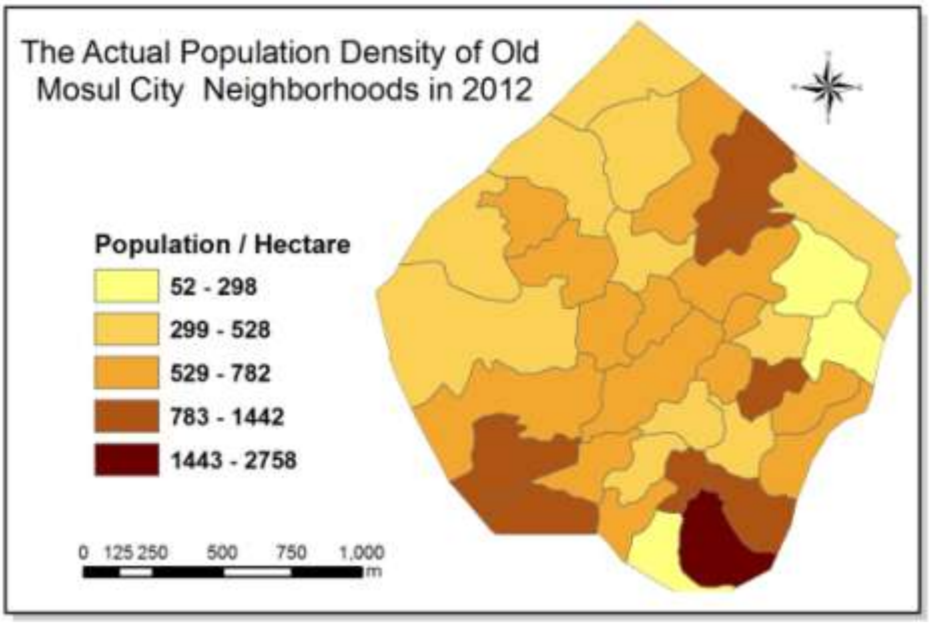
Map (3) Residential use of old Mosul city in 2012



Source: Iraq, Ministry of Municipalities, Ditto

Thus, the research succeeded in producing an effective mapping of the coloredchoroplethmap. The map is characterized by an ideal color model with high visual perception, a statistically optimized classification method, and an embodiment of high-accuracy information. (See Map 4 and Appendix 1).

Map(4): Effective Cartographic Representation of the Colored Choropleth Map



VI. CONCLUSIONS.

1- GIS users should educate themselves in basic cartography rules in order for gain a higher level of objectivity in their mapping skills.
2- The choice of the number of categories in a colored choropleth map depends on the designer's opinion in determining the appropriate number of categories that is commensurate with the number of polygonsof the mapped area. In other words, the number of categories should not be so large that it overwhelms the map reader with insignificant details and not be so that itleads to an overgeneralizedchoropleth map thatoverlooks important details.
3- Changing all threecolor characteristics; hue, value and saturation from one category to another ina GIS choropleth color map results in a higher level of visual perception of these categories.

4- The process of exploring data before representation on a colored choroplethmap anddeterminingthe frequency distribution of data, preventsGIS usersfrom falling into the problem of misleading representations that conflict with reality, and prevents the map from missing its important functionas an effective means of communication between its producer andits readers.
5- Determining the areas of actual residential use, when representing the population density data on the colored choropleth map, will eliminate inaccuracies in the design,production and interpretation of population density maps in the GIS.

REFERENCES

1.Lu,X.Z , An Object-Oriented Expert system of Quantitative Cartography, in:16th

2.ICC, Cologen , 1993, P.184.

3.(1) Arthur H. Robinson, *et al*., Element of Cartography, 6th ed., Joun Wiley & Sons, INC, 1995, P.17.

4.Borden D. Dent, Op. Cit., P.143.

5.B. Rouleau, Basic cartography for students and technicians ACI, vol.1, international cartographic association Enesco cartographague international English verision, 1987, P.99.

6.Fred,W.Billmeyer and May Saltzman, Principles of colour technology ,John wileyNew York, 1966, p.37.

7.K. Grosser, How to Acquire Reliable Rules and knowledge for Map Design expert systems (Mades) ,16th ICC, Cologen, 1993, p.1153.

8.Robert Parsons, Statistical Analysis, Second Edition. Northeastern University. New York, 1978, P.P. 57-58.

9.ESRI, ArcGIS Desktop 10.5 Software help , Data Classification.

10.Time Ormsby, *et al*., Getting to Know ArcGIS, ESRI Press, California, 2004, P.131.

11.Borden D. Dent, Op. Cit., P146.

12.Time Ormsby, *et al*., Op. Cit., p.131.

13.Borden D. Dent, Op. Cit., p.146.