1998

Issues in representing ethnic residential segregation

Wanda Teresa Gayle Smith
Wilfrid Laurier University

0-612-33809-6

Issues in Representing Ethnic Residential Segregation

by

Wanda T. G. Smith

Honours B.A., York Univerisity, 1995

THESIS

Submitted to the Department of Geography
in partial fulfillment of the requirements for
Master of Arts
Wilfrid Laurier University
1998

Abstract

This study will look at some of the issues involved in representing ethnic residential segregation. Segregation studies rely heavily upon indices and maps. However, both of these are sensitive to the spatial boundaries used. As well, maps have a visual element which affects the nature and degree of representation. Toronto has been chosen as the area of study because of its high degree of ethnic diversity. Five indices that describe the five dimensions of segregation (Massey and Denton, 1988) will be calculated at four standard levels of aggregation. As well, several types of maps will be produced to illustrate the cartographic alternatives available for representing segregation. Notably, although many of the cartographic techniques are not new, they are rarely used because they are too time-consuming and because these options are often not available in standard mapping packages.

Table of Contents

CHAPTER 3: RESULTS -- SEGREGATION INDICES

CHAPTER 4: RESULTS – CARTOGRAPHIC REPRESENTATION

CHAPTER 5: CONCLUSION

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

1.1. Introduction

Segregation is usually identified by segregation indices and proportional maps. However, both are dependent upon the level of aggregation and type of spatial unit used. Although there have been many studies on ethnic residential segregation, most have not addressed the common problems of arbitrary spatial boundaries, MAUP and areal averaging. As well, most have not addressed the changing nature of representation. Significantly, new technology and GIS have allowed us to produce maps and indices that were previously impractical or impossible to produce. For example, new technology has allowed us to use higher levels of detail and larger amounts of data. It has also allowed us to produce maps and indices more easily. The ability to readily produce many versions of the same reality can illustrate the affect of changing spatial boundaries. Thus, new technology has allowed us to examine representational issues in greater depth.

To explore these issues, this study will focus upon issues in representing and measuring ethnic residential segregation in Toronto's census metropolitan area (CMA) for 1991. Toronto has been chosen because its ethnic population is highly diverse; Toronto is the city of choice for Canadian immigrants. Both indices and maps will be produced for Toronto's primary ethnic groups. To illustrate the sensitivity of indices, this study will calculate five indices (recommended by Massey and Denton [1988] to represent five dimensions of segregation) at different levels of scale and aggregation, e.g. enumeration areas, census tracts, census subdivisions, federal electoral districts. In addition, indices dependent upon area will be calculated, with non-residential areas removed. This will be done to see how dependent these indices are upon area. Index values will be examined to see:

- if index values change significantly at different levels of scale and aggregation;
- if index rankings change at different levels of scale and aggregation;

- which spatial distributions are affected most by changing levels of scale and aggregation.

In essence, this study will illustrate the difference in values that result from changing spatial boundaries and unit size.

In addition, various maps will be produced (choropleth, dasymetric, dot distribution, proportional symbol, surface, etc.).[i] Although maps are also affected by areal averaging, MAUP and arbitrary spatial boundaries, they have additional features or concerns. Maps will therefore be looked at in terms of:

- the visual impact of arbitrary boundaries;
- the suitability of different maps to population data;
- the suitability of different maps to different levels of scale and aggregation;
- the different ways of dealing with arbitrary boundaries, MAUP and areal averaging.

It is hoped that mapping will help to complement findings, e.g. highlight placement, pattern, etc.

Chapter 1 will provide background information regarding the measurement and representation of ethnic residential segregation. Chapter 1 will consist of three main parts. Section 1.2 will provide a brief discussion of theories that relate to segregation studies in general. Section 1.3 will provide a brief overview of the main ways of representing ethnic segregation: indices and maps. Section 1.4 will discuss common issues associated with the use of census data.

---

[i] Maps will not be produced at different levels of aggregation. Rather, an appropriate aggregation level will be used for each map for illustrative purposes only.

1.2. Background and Theories

Segregation implies the lack of integration of a group into a larger and/or core society. Residential segregation is the spatial separation and consequent clustering of groups that have similar socio-economic characteristics, e.g. ethnicity (ethnic origin, mother tongue), education, income, gender, age, etc. Residential segregation, in particular, is often associated with ethnicity and/or race. "The specific patterns of ethnic concentration in terms of spatial segregation are the accumulative affects of social norms, traditions, sanctions and the strength of the group's cohesion" (Hecht, et al., 1983, p. 98). Most of the studies on ethnic residential segregation have tried to understand the internal and external mechanisms within society that result in segregation. In general, the rate of assimilation of an ethnic group depends upon the external attitudes and controls of the core and/or external ethnic groups, and the internal attitudes and controls of a specific ethnic group. An ethnic cluster can be seen to be a reflection of inner controls, e.g. the desire of individuals within a segregated group to maintain their distinctiveness and preserve cohesion. "When an ethnic cluster persists because its occupants choose to preserve it, their behaviour reflects the internal cohesiveness of the group and its desire to maintain an ensuing ethnic enclave or neighbourhood" (Fellmann, et al., 1990, p. 195). However, an ethnic cluster can also be perpetuated by discriminatory and external constraints imposed by the core and/or external ethnic groups, e.g. a ghetto is an extreme example of a cluster perpetuated by external constraints. Acculturation is the process by which an ethnic group becomes assimilated into the core society. Assimilation, however, is a long and uneven process.

> *"...full assimilation may be seen as a two-part process. Behavioural (or cultural) assimilation is the rough equivalent of acculturation: it implies integration into a common cultural life through shared experience, language, intermarriage, and a sense of history. Structural assimilation refers to the fusion of immigrant ethnics with the groups social systems and occupations of the host society. The extent of structural assimilation is frequently measured by the degree of residential segregation that sets the minority groups from the larger general community" (Fellmann, et al., 1990, pp. 180-1)*

Segregation is therefore the spatial manifestation of the lack of structural assimilation. Although there are many different forms of segregation (occupational, gender, etc.), residential segregation is the most common way of measuring the absence of structural assimilation. Residential segregation is usually represented by spatial indices (statistics) or by maps. However, the type of index or map chosen is dependent upon the definition of segregation used. The following is a quick review of major theories that deal with ethnic residential segregation.

> *"... Neither has there been a systematic review of theory bearing on ethnic segregation. Most work in the area has been conducted using a framework derived from the Chicago School of Urban Ecology. Whether researchers have been partisans or critics of this tradition, it has been the guiding beacon for 60 years of research in the field" (Massey, 1985, p. 315).*
>
> *"Reduce all social relations of space and it would be possible to apply to relations the fundamental logic of the physical sciences" (Park, in Jackson, et al. 1984, p 161).*

Ethnic residential segregation reflects larger processes of social change and economic development. Most of the studies on ethnic residential segregation have been based upon a framework derived from the Chicago School of Urban Ecology (1920s). This framework rests upon the concept of human ecology, the study of people interacting with their environment. One of the basic assumptions of ecological studies is that the same processes of plant and animal adaptation to the natural environment can be applied to human adaptation to the urban environment. Also basic to the ecological approach is the assumption that there is a direct and measurable relationship between social relations and physical distance. As Park (1926)[ii] has stated, "it is because social relations are so frequently and so inevitably correlated with spatial relations, because physical distances so frequently are, or seem to be, the indexes of social distance, that statistics have any significance whatever for sociology" (Park, in Jackson, et al.,

[ii] In the ecological tradition, Park's essay on "The Urban Community as a Spatial Pattern and a Moral Order" (1926) has been considered 'the fountainhead from which all else flows'.

1984, p. 159). Significantly, the ecological concept sees two main areas of focus. "First, it focuses on the characteristics of people in certain places and the tensions that occur at the boundaries separating people of different characteristics. ... Second, the ecological approach applies to processes that relate to groups rather than to individuals and can, therefore, be applied to aggregations of people and households that are collected by the census and published in census units (such as census tracts, blocks, and enumeration areas) of one kind or another" (Yeates, 1990, p. 149). Because of the ease of obtaining and using census tract data, the ecological approach has remained popular.

In general, the ecological approach has suggested that there is a link between urban growth and differences in ethnic and socio-economic distribution. Urban growth is therefore thought to be influenced by increased specialization and differentiation of economic activity. This, in turn, is accompanied by increased immigration. The compartmentalization that results reflects the spatial distribution of neighbourhoods according to ethnic and socio-economic features.

> *"In the classic model of urban growth, the areas of transition between the expanding business district and the retreating middle- and upper-class residential areas, provide housing opportunities for the relatively unskilled workers, including native-born internal migrants as well as immigrants from abroad. Ethnic and racial enclaves develop and serve as general reception areas to assist the new arrival in getting established. Subsequent moves from these initial areas of settlement, either by the original migrants or their children, are possible as the individuals manage to acquire the necessary skills and economic means to improve their position in the economic and social system. It has been generally assumed that those who remain in the original reception areas do so because they have not acquired the necessary skills to improve their economic status because they have experienced racial or ethnic prejudice, or because of their own preferences. In any event, the patterns of residential distribution exhibited or perceived by various ethnic or racial groups would tend to reflect the extent of their adaption to the social and economic system of the community as a whole" (Breton, et al., 1990, pp. 11-12).*

The traditional model therefore assumes that in time immigrants will eventually move from the inner core (the initial areas of reception) to the suburbs (the periphery). It also assumes that immigrants and/or their offspring will eventually become integrated into the dominant or core society. Segregation is therefore thought to decline with each new generation. Thus, in ecological terms, segregation can be seen to be a 'natural'[iii] process. Invasion/succession -- a concept central to the ecological model -- refers to a process where as one group moves into an area, another group moves out. Invasion/succession thus suggests that change occurs in a cyclical fashion. The driving forces behind succession are immigration and social mobility, e.g. newer immigrants move into areas that are vacated by socially mobile classes.

Although the traditional model is still widely used, it has also been criticized. In general, the ecological model has been criticized for not being complex enough to address the degree of ethnic diversity that is found in most North American cities today. That is to say that immigrant settlement patterns in urban America are often more complex and varied (Ray, 1994). This is especially true of settlement patterns in Toronto.

> *"... immigrant settlement patterns in Toronto are more diffuse and complicated than the traditional invasion/succession model would suggest" (Ray, 1994, p. 264).*
> *"... [the] invasion/succession model traditionally used to explain the entry of immigrants into the housing market, and the social geography of immigrant groups in the city (Ward, 1989) provides relatively little insight into the dynamics of immigration settlement in contemporary Toronto" (Ray, 1994, p. 262).*

Also, although the traditional model was useful for Chicago at the time, the basic assumptions[iv] of the model have become increasingly problematic. For example, since the 1920s there has been more public intervention, more than one urban centre/nuclei, and a growing awareness that ideal conditions do not exist (places are not perfect isotropic plains). Furthermore, the traditional model has been criticized for

---

[iii] 'Natural' implies unplanned processes, e.g. residents of urban areas 'naturally' compete for the most desirable areas to live in accordance with socio-economic demands.

[iv] The five basic assumptions of the traditional ecological model are that spatial expansion is dependent upon (a) economic growth, (2) ethnic diversity, (3) class stratification, (4) a single centre, (5) an isotropic plain.

being economically deterministic (Ley, 1993). In addition, the assumption of full assimilation through time has been questioned, e.g. many ethnic groups have resisted assimilation and taken on a degree of permanency.

By the 1950s, the ecological model was revised to include social area analysis and later factorial ecology.

> *"The simplest and most widely accepted ecological theory for explaining urban residential structure is that of social area analysis (Shevky and Bell, 1955) and its offspring factorial ecology (Berry and Horton, 1970; Timms. 1971; Berry and Kasarda, 1977). It holds that social and residential differentiation are a function of the level of societal development as societies grow economically, their inhabitants become increasingly heterogeneous with respect to three fundamental dimensions: socio-economic status, family structure, and ethnic background. Over time, people are increasingly differentiated by occupation, income, and education, creating a complex, multileveled class structure. Moreover, the shift of economic production out of household into factory spurs rising female employment, and generates a greater variety of family types. Finally, sustained economic growth creates a strong demand for labour, much of it unskilled, which is met through ethnically diverse immigration (Kindelberger, 1967; Piore, 1979). These societal developments are, in turn, reflected spatially" (Massey, 1985, p. 316).*

Despite advantages, however, social area analysis has received much criticism. Specifically, criticism has centered around the limitations of a three-dimensional view of social development, e.g. spatial indices were used to classify areas along three dimensions. Although social area analysis was replaced in the 1960s by the more inductive factorial ecology, it has gained importance because of its influence upon and stimulus of further research in urban studies. Significantly, factorial ecology uses factor analysis or principal component analysis to classify socio-economic, demographic and housing data into different components. Factorial ecology refers to "the application of either factor analysis or principal component analysis to matrices of socio-economic, demographic, and housing data for small intra-urban districts (census tracts) to test the hypothesis that the pattern of residential differentiation (segregation) can be accounted for by a small number of general constructs" (Johnson, in Gregory, et al., 1994, p. 186). Therefore, constructs or dimensions emerged from the data rather than having predetermined dimensions.

The move to social area analysis and later factorial ecology reflected the general shift within geography itself. Significantly, in the 1960s, the empirical (descriptive) gave way to the positivist (analytical). The spatial/positivist school was favoured because it was considered to be free of the cruder biological qualities of human ecology. The emphasis of spatial analysis is upon pattern, inference and hypothesis testing. "The language of spatial analysis accommodates variables which may be specified precisely and are therefore objective and measurable" (Ley, 1993, p. 4). Yet reliance upon the objective, linear, and measurable is restrictive in that it disqualifies variables that do not fit the specific language of spatial science. Notably, the advancements and greater accessibility of computers and GIS in the 1980s and 1990s, have strengthened the emphasis upon spatial science. Spatial science, however, is ideally accompanied by a greater awareness of the limitations of the positivist approach.

> *"The application of structuralism in modern geography has been correspondingly diverse, but in general the concern with the intellectual exposure of structures has been retained. Structuralist geographers have recognized the importance of clarifying the theoretical status of the constructs deployed in empirical work and of penetrating the outward forms of social life; while they have tended to resist the linguistic turn -- the view that all social life is 'like a language' -- and thereby have defined the boundaries of the symbolic domain more carefully" (Gregory and Smith, 1986, p. 462).*

The structuralist approach introduced in the 1970s has also had an influence upon segregation studies. Although the structuralist tradition is extremely broad,[v] its branches have revealed common elements and shared deviations from the ecological. The structuralist approach differs from human ecology in that it eliminates cultural factors. "Cultural factors are not included which is intentional, and

---

[v] Marxist theories have been an important part of the structuralist tradition within geography. Influential in this regard has been the work of D. Harvey (e.g. 'Social Justice in the City', 1973).

which corresponds to the focus ... on social structures rather than on culture." (Blau, 1977, p. 159). Thus structuralism places greater emphasis upon definition and theoretical background: studies are not thought to be value-free but to have implicit meaning. From a structuralist point-of-view, segregation can be seen to be influenced by social structures rather than by natural/cultural forces. Segregation is no longer a natural/cultural phenomenon but evidence in itself of the existence of inequality. Segregation reflects as well as perpetuates unequal structures within society. Empirical data such as indices and maps are still commonly used, however, they play a secondary role to theory. In the structuralist tradition, empirical detail is used to support the conclusions of abstract and theoretical research (Jackson et al., 1984). Therefore, some forms of structuralism can be considered more abstract than human ecology.

Significantly, the structuralist tradition falls under the more inclusive category of macrosociology. "The macrosociological approach is concerned with the discovery of structures within the society as a whole, the examination of large scale relationships in society, of the relationships among the structures within the society. Structures set the tone for behaviour, the context within which behaviour takes place" (Anderson, 1996, p. 1). A macrosociological theory is a deductive theory of social structure which "seeks to explain relations among various parts of entire societies in terms of differentiation of [various] parts" (Blau, 1977, p. 2). Macrosociological theory is based upon the premise of parameters (attributes that differentiate social position) (Blau, 1977). Parameters are divided into two subgroups, nominal and graduated. Nominal parameters include sex, race, religion, ethnic affiliation, occupation, place of work, national origin; while graduated parameters include education, income, wealth, prestige, power, intelligence. Nominal differentiation is called horizontal differentiation or heterogeneity. Graduated differentiation is called vertical differentiation or inequality. Basically, the macrosociological approach suggests that the higher the degree of heterogeneity, the lower the degree of discrimination and inequality. However, although a society may be heterogeneous, this can be counteracted by several forces, one of these being segregation. In general, the macrosociological approach suggests that opportunities for social contact are diminished by physical distance.

> *"Physical barriers -- be they prison bars or oceans or long distances -- naturally impede social associations. Groups that are located far apart have few opportunities for social contacts. Regardless of the great heterogeneity of a society, the chances that persons from different groups meeting are small if different groups live and work in different places. The spatial segregation of a group limits the influence of heterogeneity on intergroup relations. It does not completely obliterate this influence, however, particularly in contemporary societies where modern means of transportation and extensive mobility help overcome the obstacles of physical distance to social associations" (Blau, 1977, p. 90).*

The structuralist approach, however, has also met with criticism. For example, this approach has been criticized for being structurally deterministic and denying human agency. "Men and women are reduced to the passive 'bearers' of a vector of structural determinism" (Gregory, et al, 1986, p. 462). Also, the structuralist approach has been criticized for lacking a cultural and historical perspective.

By the 1980s, however, other approaches began to surface within geographic studies, e.g. humanism, realism, postmodernism, etc. These approaches criticized previous approaches such as empiricism, positivism and structuralism. The newer approaches placed greater emphasis upon historical perspective, human agency, and the ethnographic method. However, newer models have not replaced earlier models in terms of measurement and structural segregation. Hybrid models such as the ecological [under pragmatism] (Smith and Jackson, 1984), and the core-periphery (Hecht, et al, 1983), have been recommended for current segregation studies. Despite criticism, segregation indices and other quantitative measures are still commonly used in segregation studies today. A "segregation index ... always fails to cope with the spatial element because it is unable to identify specific areas of ethnic segregation or concentrations across space or to exhibit a spatial population distribution" (Hecht, et al, 1983, p. 155). Maps complement indices because they are able to show what indices are not. Also, "analyses [involving indices] will tell, at best, a portion of a complex story and, at worst, may inadvertently contribute to the prejudice. Researchers reproduce ethnocentric and racist bias when they portray ethnocultural groups as 'naturally' homogeneous (a pitfall difficult to avoid in the context of census data), 'naturally' separate, and 'naturally' antagonistic to one another. Recent statements on the social construction of ethnic and racial categories offer an important corrective in this regard" (Hiebert, 1994, p. 258). Thus geographers today are more aware that indices and maps should not be viewed in isolation. Despite this, however, the traditional models/methods are still commonly used to measure and represent ethnic residential segregation (e.g. the question of 'why' is addressed more fully by newer theories/methods, but the question of 'what' still relies upon the traditional). Specifically, the re-emergence of quantitative methods has been facilitated by the growing use of computers, GIS and computer mapping. The use of computers has allowed the use of new procedures and maps that were previously not possible, e.g. certain indices are now used that were previously too complex and time consuming. Also, the growing importance of GIS and digital representation has placed greater emphasis upon issues of representation.

1.3. Measurement and Representation

All representation is, in essence, an abstraction. Representation itself is the result of a process called signification. Signification is the process of using signifiers to construct meaning. A signifier is thus a medium that is used to produce meaning, e.g. a cross could be a signifier for a church. For maps, the process of abstraction into signifiers produces a 'representation of reality'. Although traditional maps with crude signifiers are still used in cartography, more realistic representations of geographic phenomena are becoming increasingly common, e.g. remote sensed images and glossy GIS maps. These realistic images have been enhanced and promoted by the explosion of electronic technology, e.g. high resolution remote sensing, advanced GIS (including DEMs, virtual GIS, animation), etc. Notably, the image in geographic representation is now moving closer to the realistic ideal of forgetting the process of signification.

Although there are many levels of abstraction, the two main stages include iconic and symbolic. Iconic tries to "directly portray certain visual aspects of the territory in question" (e.g cartographic representation, aerial photography, etc.), while symbolic "utilizes purely conventional signs and symbols, like letters, numbers, and graphic devices" (e.g. indices, mathematical equations, etc.) (Turnbull, 1993, p. 3). At every stage of abstraction, however, information is necessarily lost. That is to say, that at each stage of abstraction, representation becomes more general and abstract. Representations that are more abstract are clearer and show relationships that might otherwise be lost. However, they are also open to manipulation and further removed from the real world.

This study will focus upon the two main ways of representing ethnic segregation: (1) segregation indices and (2) maps.

1.3.1. Segregation Indices

The most common way of representing or measuring segregation is through a segregation index. Segregation indices are descriptive statistics that describe the spatial distribution of variables. Familiar descriptive statistics include mean, median, standard deviation, z-score, skewness, etc. Descriptive spatial statistics that are less familiar include mean centre, shape index, standard distance, centroid computations, autocorrelation, etc. The use of standardized spatial indices have allowed different variables and different studies to be easily compared. The most common segregation index is the index of dissimilarity. The index of dissimilarity is a measure of the relative concentration of a group by spatial unit, compared to its relative proportion to the city as a whole. From 1955 to 1976 the index of dissimilarity was used almost exclusively. Influential in this regard was a study by Duncan and Duncan (1955) that suggested that the index of dissimilarity was the obvious index of choice because its data requirements were minimal and its computation simple. Also, the index of dissimilarity was advocated because it minimized population composition and group size.

> *"Duncan and Duncan (1955) ushered in a long era of peace by demonstrating that there was little information in any of the prevailing indices not contained in the index of dissimilarity and the minority proportions. ..... For more than 20 years afterwards, the dissimilarity index served as the standard segregation measure, routinely employed to measure spatial segregation between social groups" (Massey and Denton, 1988, p. 281).*

In addition, the dissimilarity index has the advantage of being able to be viewed graphically in the form of a Lorenz curve. The Lorenz curve is a plot of the cumulative proportion of group x against the cumulative proportion of group y (variables are sorted in descending order). Several indices have been derived from the Lorenz curve. One of these is the dissimilarity index. The dissimilarity index represents the maximum distance from the Lorenz curve (the diagonal line).

However, since the mid-1970s, the dissimilarity index has met with disfavour. Influential in this regard was a pivotal study by Cortese, Falk & Cohen (1976). After this study, controversy arose as to whether this index was the best index to use. By the mid to late 1980s, segregation studies were "in a state of theoretical methodological disarray" (Massey and Denton, 1988, p. 282). The period between 1976 and the late 1980s is now referred to in segregation studies as the period of indice wars. "One reason for the disagreement is the absence of a clear set of criteria, derived from a comprehensive definition of segregation, which can be used to evaluate the various measures that have been proposed" (James and Tauber, 1985, p. 2). Although comparability of research suffered because of the use of many indices, debate was useful in that it focussed attention upon the relationship between the definition of segregation and the indice chosen. Also, debate was useful in that it raised questions regarding the appropriateness of various indices. Common to this period was the notion that a single index was not sufficient to address and/or represent segregation. Therefore researchers recommended different indices to represent different aspects and operational definitions of segregation (see James & Taeuber, 1985; White, 1983; Stearns & Logan, 1986; Lieberson & Carter, 1988). This coincided with broader developments within segregation studies in general. These developments included the realization that the different facets of segregation resulted in different perceptions and/or experiences. An important study during this period was a study by Massey and Denton (1988). Massey and Denton argued that the index of dissimilarity was limited in that it only measured 'evenness', one of five dimensions that made up spatial segregation, e.g. evenness, exposure, concentration, centralization and clustering. These dimensions, they suggested, were conceptually distinct even though their outcomes were often similar. Their study compared twenty segregation indices on sixty CMAs in the United States at the census tract level. With the help of factor analysis, they empirically supported a five-dimensional view of segregation, and recommended five indices to help describe them, e.g. (1) Index of Dissimilarity <Evenness>, (2) Interaction Index <Exposure>, (3) Relative Concentration Index <Concentration>, (4) Absolute Centralization Index <Centralization>, and (5) Spacial Proximity Index <Clustering> (see Chapter 2).

> *"Urban spatial structure is inherently multidimensional (Timms, 1977), and residential segregation, in particular, does not stem from a single process, but from a complex interplay of many different social and economic processes that generate various constellations of outcomes interpreted as 'segregation.' These spatial outcomes are created through the combination of five basic distributional characteristics. Evenness is the degree to which groups are distributed proportionately across areal units in a city. Exposure is the extent to which members of different groups share common residential areas within a city. Concentration refers to the degree of a group's agglomeration in urban space. Centralization is the extent to which group members reside toward the centre of an urban area; and clustering measures the degree to which minority areas are located adjacent to one another" (Massey and Denton, 1988, p. 310).*

Although other studies have recommended other dimensions and indices, the five dimensions have been the most well received. Hypersegregation, a recent concept, is based upon the five-dimensional view of segregation. Hypersegregation refers to high index values (above 0.6) of at least four of the five dimensions (Denton, 1994). "Hypersegregation suggests ... a pyramiding at the spatial level, implying drastic isolation and describes the difficulty of escaping the effects of multidimensional layers of segregation piled on top of one another" (Denton, 1994, p. 50). Although the five indices are commonly used in segregation studies today, the index of dissimilarity is widely used in isolation. This can be attributed to the fact that the index of dissimilarity allows results to be easily compared through time. Also, this index is easier to compute and has fewer data requirements.

1.3.2. Cartographic Representation

Maps are also widely used in the representation of ethnic residential segregation. Maps complement indices by showing pattern, direction, and place. These can be further distinguished by frequency, size (magnitude) and shape. Common distribution patterns include uniform, random and clustered. Maps can show directional information by showing where groups are moving to or moving from. Place refers not only to location but also to more meaningful/recognizable areas (e.g. Chinatown). Recognizable areas can also refer to neighbourhoods with which people can readily identify.

The question of what makes an optimal map has created much debate. Ideally the map maker's perception and the map user's interpretation are the same. However, this is not always the case. Different conceptual models have been developed to try to capture the various stages of the map making process. Most models appear to follow a continuum from the map maker's message (or facts that he/she wishes to convey) to the map user's interpretation of the final map. Maps consist of subjective interpretation and selection and objective grounding in the real world. "Maps represent an abstract view of some portion of the world with an emphasis on selected features" (Dent, 1993, p. 4). Since maps are, in essence, an extraction of selected features from reality, they represent a simplified, partial version of the truth. According to Dent (1993), "the transformation process [from data in reality to data in a map] is a process of simplification and symbolization" (Dent, 1993, p. 13). Cartographic abstraction (or generalization) can be broken down into four main areas: selection, classification, simplification and symbolization. Selection involves the choice of scale, projection, variables, etc. (Dent, 1993) Classification refers to the process of placing phenomena into similar groups for the purpose of clarification. Simplification involves the removal of unwanted detail in order to make the message more clear. Symbolization refers to the process of matching real world phenomena to appropriate symbols.

Notably, the cartographic act of abstraction -- subjective and exclusive -- is open to distortion and misrepresentation.

Part of the selection or simplification process involves deciding which level of measurement to use, e.g. nominal, ordinal, interval, or ratio. "Data can be mapped as categorical (nominal or ordinal) or numerical (discrete or continuous)" (MacEachren, 1994, p. 13). The different levels of measurement can be represented by four main object classes (or spatial dimensions): points, lines, areas, and volumes/surfaces. Eight graphic variables that can be used when representing the object classes are location, size, colour (hue, saturation, value), orientation, shape, arrangement, texture, and focus. Most map symbolization is built upon these variables. "A good map designer knows how to match graphic variables with these spatial dimensions and measure levels in effective ways" (MacEachren, 1994, p. 13). Mapping is thus a complex task of extracting and presenting information. Although certain rules should be followed, one should always keep in mind the purpose of the map and the user. Aesthetic appearance is also important; information should be presented in an effective and aesthetically pleasing manner.

In order to represent segregation, it is necessary to map population data. Population data can be presented in all object classes: points, lines (flow of people), areas and surfaces. Ideally, the choice of object class depends upon the type of data represented. According to MacEachren (1994), data ranges from discrete to continuous, and abrupt to smooth. Population data is considered to be both smooth and discrete. Despite this, the areal representation is the most common representation. The areal representation (choropleth map) has the advantage of portraying the way that data is collected, able to show density, and being readily available in standard mapping software. However, choropleth maps have the disadvantage of giving visual emphasis to unpopulated areas. Dot distribution maps are considered more appropriate representations of census data because they are of the same object class [smooth and discrete] (MacEachren, 1994). However, dot distributions can be very time consuming and there is no way of knowing the distribution of already aggregated data. "Dot maps are very time consuming to construct and require considerable research" (Dent, 1993, p. 158). Notably, research into the placement of dots can be very time consuming. Proportional point symbols are also used frequently. Proportional point symbols have the advantage of being aesthetically pleasing, easily produced, and able to show differentiation of intensity. An alternative is the surface representation. A surface representation shows population as a volume. Although surface representations come in several formats, the most common is the raster format, e.g. raster/grid/mesh maps. The surface model offers the advantage of being aesthetically pleasing and showing gradual transition. However, surfaces are open to manipulation. Also, some researchers have questioned the validity of converting arbitrary areal aggregates to surfaces, or of showing population as being continuous. Other alternatives include cartograms and statistical maps[vi]. These maps offer the advantage of clarifying the cartographic message by removing unwanted information. However, like surface models, these maps are open to manipulation. Also, they can be confusing to the reader. Notably, the choice of map depends upon the variable chosen, the detail required, and the message the author is trying to convey.

[vi] Statistical maps in this study refer to maps that use higher order equations instead of absolute and ratio values.

1.4. Census Data

This section will focus upon (a) the nature of census data, and (b) the definition of ethnicity.

1.4.1. The Nature of Census Data

All representations are limited by their input data. Since most maps and indices use census data, they are limited by the properties of the census data itself. Most of the problems associated with census data are related to way that it is collected. Census data, which fall under the areal category of 'arbitrary collection units' (Chrisman, 1989), are areal units that have previously imposed boundaries, e.g. enumeration areas (EA), census tract areas (CT), etc. In other words, "the positional description of the object precedes any attributes assigned. These maps are choropleth maps in the purest sense because the places exist, then they are filled" (Chrisman, 1989, p. 23). Socio-economic data is associated with this type of data because it is usually collected by enumeration, and includes such things as county boundaries, national boundaries, electoral boundaries, etc. Notably, there are different criteria for creating census spatial units. The fixed boundaries for a census year are used as collection units for all variables collected by Census Canada. Also, neighbourhoods are not static entities, e.g. the socio-economic characteristics of neighbourhoods often change. Arbitrary boundaries are in contrast to the preferred natural boundaries found in categorical coverages[vii] (Chrisman, 1989). In a categorical coverage, "some system of classification (soil taxonomy, vegetation classes, and even the list of taxable parcels) logically precedes the map" (Chrisman, 1989, p. 24). The assumptions are that the underlying

---

[vii] Well-known examples of categorical coverages include soil class maps, landuse maps, forest cover maps, climate zone maps, etc.

phenomena are continuous and spatial units are adjusted to show categorical distinction. Since this logic does not apply to arbitrary collection units, certain issues are raised. Some of these issues include:

a) representational issues;

b) heterogeneity within the collection unit;

c) modifiable areal unit problem, and

d) aggregation and scale effects.[viii]

1.4.1.1. Representational Issues

One problem with census data is that the phenomena being mapped are aggregated individuals. Although individuals are single entities, they are usually presented in different object classes such as areas or surfaces. However, it can be argued that the viewing of population as individual, areal neighbourhoods or continuous density functions is really a function of scale. Uncertainty and error can result when one type of structure or object class is transformed into another. MacEachren (1994) calls this 'seeing wrong' (type I error) and uses the example of transforming a 3-d surface into a 2-d plane. The same can be said about transforming count or population data (single entities) into areal units or continuous surfaces. If this is not done carefully, it can result in different interpretations and/or misleading results. These aspects are further complicated by scale and aggregation effects. Furthermore, in GIS, "because the system representation of the population-based data for the census district is no longer of the same spatial class as the real world phenomenon it represents, both mapping and manipulation operations are severely restricted" (Martin, 1991, p. 61).

[viii] In addition, although much of the information is collected at an interval level, much is portrayed at an ordinal or nominal level in order to clarify the message or pattern. For example, the choice of a classification taxonomy can create very different perceptions.

1.4.1.2. Heterogeneity Within the Collection Unit

Heterogeneity within the collection unit questions the assumption of the areal unit. The basic assumption behind the areal unit is that the phenomenon being represented is homogeneous. Since this is rarely the case with socio-economic phenomena, problems of interpretation can occur. Also, heterogeneity within collection units can produce misleading visual impressions (e.g. a spatial unit can be visually mistaken for being evenly distributed). This is a general problem with all choropleth maps. A dasymetric map is a type of choropleth map that tries to account for this. "The dasymetric technique recognizes that what is being mapped varies in intensity within the mapping units. To accommodate this, each mapping unit is divided into a number of smaller units... Each of these subdivisions is defined so that it is relatively homogeneous in terms of the phenomenon being mapped" (Campbell, 1991, p. 218). However, once aggregated, the level of detail that is lost cannot be retrieved. Therefore, creating a true dasymetric map is nearly impossible. Yet dasymetric techniques (the removal of non-residential areas) have been useful in showing some variation within the collection unit. Notably, Statistics Canada uses the term ecumene population (a Greek word meaning inhabited land) for this type of dasymetric representation.

1.4.1.3. Modifiable Areal Unit Problem

The Modifiable Areal Unit Problem - MAUP (Openshaw, 1984a) is related to the modifiable nature of spatial collection units. This refers to the problem of the different orientation of boundaries. In fact, there is an unlimited number of possibilities for allocating different sets of individuals to different zoning units. Consequently, results are dependent upon the type of zone that is used. Numerous suggestions have been made to try to address this problem (see Openshaw, 1984; Openshaw, 1989; Hunt, 1993; Dorling, 1995). "The simplest is merely to illustrate it by using multiple boundaries, and with interactive visualization it is possible to redraw images instantly using different boundaries to see the effects of these choices" (Dorling, 1993, p. 177). Another simple choice involves aggregating to higher more meaningful spatial units -- although still remaining arbitrary. The MAUP is comprised of two problems, scale and aggregation. These problems also include arbitrary choices. The problem of scale focuses upon the question of how many zones should be used, e.g. what level of aggregation should be used. The problem of aggregation focuses upon the question of which zone should be chosen at a given level of aggregation (Martin, 1996). However, MAUP differs from aggregation in that the reduction in variance is not uniform (Barber, 1988). This raises the question of validity when comparing different areal units; for example, comparing postal code districts with enumeration areas.

1.4.1.4. Aggregation and Scale Effects

This refers to differences of interpretation that occur when looking at the same data at different levels of scale or aggregation. For example, if we compare data at the enumeration level to the same data at the census subdivision level, the mean for the entire study area would likely stay the same, while the variance would likely decrease. It is important to note that once data is aggregated, the level of detail that is lost cannot be retrieved. With scale, the choice of either a neighbourhood or a provincial level of study affects the final outcome. The choice of scale is usually dictated by the purpose of the study.

Segregation studies generally use a metropolitan level of scale and a neighbourhood (census tract) level of aggregation. This is especially true for ecologically-based studies. However, other levels of aggregation have also been used. For example, Smith (1989) has advocated the use of data at the enumeration district level within Britain (enumeration area equivalent). This is under the assumption that "segregation in British cities is often intense, albeit on the scale of enumeration districts or streets. That the scale is small ... is a reflection of the comparatively small size of the coloured minority populations [in Britain] rather than of any particularly enlightened community attitude towards interracial housing" (Smith, 1989, p. 38). However, Weiher (1991) has supported the use of municipality level aggregation for studies in the United States, e.g. "segregation by municipality in suburbs to be compared with segregation by neighbourhoods in the central city" (Weiher, 1991, p. 46). Furthermore, he has suggested that although distinctive neighbourhoods exist within a city, their boundaries are hard to define. "Political boundaries are more visual than neighbourhood boundaries, and information about politically defined residential space is likely to be acquired at less cost than information about neighbourhoods" (Weiher, 1991, p. 45). In addition, residential segregation is either promoted or discouraged by different legislation that accompanies different political boundaries. Thus, changes to the level of scale and aggregation changes the degree of segregation perceived. A methodological study at different levels of scale and aggregation is therefore important because it influences our perception of the degree of ethnic segregation that exists.

1.4.2. Definition of Ethnicity

A definition of ethnicity is both ambiguous and multi-faceted. That is to say, there are many ways of defining ethnicity. For example, features or markers of ethnic identity ascribed by Rotherchild (1981) include:

- "race, as a phenotypical feature;
- kinship, through blood ties or alleged common ancestry;
- language, as a vehicle of communication or a symbol of identity;
- religion, as a type of social allegiance;
- a customary livelihood;
- a strong territorial identity; and
- historical political autonomy" (Hecht, et al, 1983, p. 3).

Markers such as these can be used as inclusive or exclusive features to define ethnic identity. An ethnic group is therefore a group to which one identifies or to which one can be identified according to selected features/markers. However, ethnic identity is often more complex and varied. For example, definition can be complicated by the fact that people often have multiple cultural links. Also, generational issues are often a problem; a third or fourth generation Canadian may have either a weak or strong attachment to his/her ethnicity. Statistics Canada has different ways of defining ethnic identity, e.g. ethnic origin (20% sample -- single and multiple responses), mother tongue (100% sample). Notably, the mother tongue category (100% sample) is considered to be more accurate than the ethnic origin category, because it is a 100% sample and thus free of sampling error. However, the ethnic origin category is more inclusive.

Because this study relies upon census data, the definition used will necessarily focus upon markers ascribed by Census Canada. In their 1991 survey, the question that was asked was:

"To what ethnic or cultural group(s) did this person's ancestors belong?"

Thus, in this context, an operational definition is limited to parameters set by Census Canada. Ethnic identity can therefore be said to reflect ethnic ancestry. Ethnic ancestry does not, however, reflect the respondent's nationality, citizenship or ethnic identity. This has changed from the 1986 Census where equal emphasis was placed upon the ethnicity of the respondent and his/her ancestors. The question asked in the 1986 Census was: "To which ethnic or cultural group(s) do you or did your ancestors belong?" The 1991 Census has thus narrowed its focus to avoid the confusion produced by a dual focus. The current question is also an improvement over earlier questions. For example, prior to the 1981 Census only paternal ancestry was considered. Also, prior to the 1981 Census, allowance was only made for a single ethnic origin. Since 1981, however, provision has been made for the inclusion of multiple origin ethnics.

1.5. Summary

Chapter 1 has provided background information and discussed theories that relate to segregation studies in general. It has also discussed some of the main issues concerning the use of segregation indices and maps. As mentioned previously, this study will produce indices and maps for the city of Toronto for the 1991 census. Chapter 2 will illustrate the methods that were used for producing segregation indices and maps. Chapter 3 will discuss the results of using segregation indices. Chapter 4 will discuss the results of using maps. The final chapter, Chapter 5, will review some of the main findings, discuss their implications, and give suggestions for future research.

CHAPTER 2

METHODOLOGY

2.I. Area of Study

The area of study is Toronto's Census Metropolitan area (CMA). Toronto has been chosen because of its high degree of ethnic diversity. In fact, the number, variety and complexity of Toronto's ethnic population has grown significantly in recent years. This has been the result of high levels of immigration. "Toronto is the pre-eminent destination for new immigrants, more of whom plan to locate in this one city than any other city or province in the country. Among immigrants coming to Canada in 1991, 27.7% intended to locate in Metropolitan Toronto (Employment and Immigration Canada 1992). In 1991, 37.7% of Toronto's CMA's population was foreign born. Changing the scale in Metro Toronto, the foreign-born population rises to just under 41%" (Ray, 1994, p. 262). Significantly, the nature/origin of the new immigrant has changed considerably in recent years. Far from the strong British influence of pre-WW2, Toronto now reflects a broad ethnic mix. In fact, the rise of visible minorities has become a important feature of the new ethnic mosaic. The change to a more varied and cosmopolitan mix has been prompted by changes to Canada's immigration polices. Prior to the 1960s, Canada tried to maintain a predominantly white society by selective advertising and by giving preference to white immigrants, especially those of European descent. However, a change in 1967 to the point system placed more emphasis upon economics and need. Thus Toronto has evolved as "the worlds most multicultural city" (Fellman, et al, p. 181). It therefore lends itself well to a study of measurement and representation of ethnic residential segregation.

2.2. Ethnic Variable

The ethnic origin variable from the 1991 Population Census of Canada has been chosen for the ethnic variable. This variable has been used because it is the most inclusive. Notably, this variable is a 20% sample that includes single and multiple origins. The population of Toronto's CMA is 3,891,265; of this, 2,920,430 are of single origin, and 937,805 are of multiple origin. The following ethnic groups have been chosen for this study because they have large populations and varying spatial patterns: English, Italian, Chinese, East Indian, Portuguese, Black, Jewish, and German. Historically the Germans have shown high degrees of assimilation, e.g. dispersed spatial patterns and low levels of segregation. On the other hand, Italians, Jews and Portuguese have tended to cluster in larger neighbourhoods and thus have exhibited higher levels of segregation (Hecht et al., 1983; Breton et al., 1990). The Blacks and Caribbeans have exhibited moderate to high levels of segregation, however, their neighbourhoods have usually been smaller and more scattered (Ray, 1994).

Also used were ethnic origins that were grouped together according to region and/or subcontinent (according to Census Canada classifications) -- see Table 2. The grouping of nationalities into major racial/ethnic groups was felt to be useful because all single origin groups could be included. As well, smaller groups could be included to demonstrate the effect of group size. In addition, the inclusion of major regional ethnic/racial groups can be useful for revealing regional spatial patterns. Significantly, the qualities and historic patterns of various ethnic groups provides an important background for comparative studies. Therefore, 21 groups have been used in this study. These include both large single ethnic groups and regional racial/ethnic groupings.

Table 2.1: Classification of Ethnic Groups by Region or Sub-continent, for Toronto's CMA, 1991

Regional Definition	National Ethnic Groups included	Population	Proportion of Total Pop.
Aboriginal	Aboriginal	5,935	.0015
Arab	Lebanese	8,380	.0022
Northern European	Norwegian, Swedish, Danish, Finnish	12,815	.0033
French	French	50,565	.0130
Western European	German, Dutch (Netherlands)	98,040	.0252
Jewish	Jewish	113,940	.0293
Black	Black	124,560	.0320
Eastern European	Hungarian, Ukrainian, Polish	134,715	.0346
South Asian	East Indian	140,340	.0361
Canadian	Canadian	264,280	.0679
South East Asian	Japanese, Korean, Vietnamese, Filipino, Chinese	352,245	.0905
Southern European	Yugoslavian, Croatian, Spanish, Greek, Portuguese, Italian	556,855	.1431
British	English, Scottish, Irish, Other British	740,430	.1903
Other	n/a	287,980	.0740
Multiple Origins	n/a	935,940	.2405
Total Population	n/a	3,891,265	1.0000

The following provides background information regarding the spatial distribution of the ethnic groups used.[i]

[i] Unless citied otherwise, the sources for background information have come primarily from Hecht et al, 1983 and Breton, et al., 1990. Also, statistical references refer to 1991 Census information.

Aboriginals

The North American Indian group makes up only 0.15% of Toronto's single origin population. Spatially this group appears to be lightly scattered throughout the metropolitan region. However, more aboriginals of single origin live outside Toronto and Mississauga.

Arabs (Lebanese)

The Arab group represents a small percentage of Toronto's single origin population (0.2%). Because of the low population and recent immigration of this group, little has been written about their spatial pattern. However, people of Lebanese descent have higher representations in North York, Mississauga and Scarborough.

Blacks/Afro-Caribbeans

The black group (African and Caribbean) reveals a low rate of assimilation and intermarriage. However, most blacks have English as a first language. This group represents 3.2% of Toronto's single origin population. Significantly, the immigration of the black group has been fairly recent (post 1970s). Although the segregation rate of this group is fairly high, it is not as high as the black segregation rates found in most American cities. In Toronto, there is a high proportion of blacks living in high density areas (e.g. highrise rental buildings) that are moderately dispersed throughout the city. In contrast to the pattern in America, the black group in Toronto is strongly represented in suburban areas. For example, North York and Scarborough have twice as many people of single black origin than the (former) city of Toronto. However, they have little representation outside Metropolitan Toronto.

British (English, Scottish, Irish, Other British)

The British group is the primary and largest ethnic group within Toronto's CMA with 19% of the single origin population. According to the 1991 census, over one million people (25%) are of single or multiple British origin. The British group has typically a high rate of intermarriage. Although this group is evenly scattered throughout Toronto's CMA, it has a higher proportion outside Metropolitan Toronto.

Eastern Europeans (Hungarian, Ukrainian, Polish)

The Eastern European group represents 3.5% of Toronto's single origin population. Much of the immigration of this group occurred after World War II. This group is moderately dispersed throughout the metropolitan region. Overall, this group exhibits an intermediate level of segregation.

French

The French is a relatively small group within Toronto's CMA (e.g. 1.3% - single origin). The French are similar to the Western Europeans in that they are geographically dispersed and have a high rate of intermarriage. Notably, there are significantly more French of multiple than of single origin, e.g. there are 174,250 of multiple origin and 49,940 of single origin, totalling 5.8% of the population.

Jewish

The Jewish population represents 2.9% of Toronto's single origin population. The Jewish group has low rates of assimilation and intermarriage. They also have a high rate of English as a first language and a high degree of economic success. Significantly, the Jewish group measures the highest level of segregation, a distinction that they have maintained for decades (Hecht et al, 1983). "The population of Jewish origin has continued to exhibit the highest levels of residential segregation following the 1932 census" (Breton, et al, 1990, p. 95). The Jewish segment is mainly concentrated along the Bathurst Street corridor, in North York (primarily), Toronto and York.

Northern Europeans (Norwegian, Swedish, Danish, Finnish)

The Northern European group is similar to the Western European group in that it shows a high rate of integration, intermarriage and English as a second language. Multiple origin responses are significantly greater than single origin responses (Halli, et al, 1990). Although this group represents only 0.3% of the single origin population, this may be misleading. The immigration period for most Northern Europeans was prior to World War II; most immigrated to Canada in the 1920s. Overall, the Northern European group is fairly evenly dispersed throughout the city.

South Asians (East Indians)

The East Indian population represents 3.6% of the single origin population in Toronto's CMA. The East Indians are also newer immigrants, with most arriving after the 1970s. The highest East Indian population is found in Mississauga, but the highest proportion is found in Brampton.

South East Asians (Japanese, Korean, Vietnamese, Filipino, Chinese)

The South East Asian population represents 9% of Toronto's single origin population. This group has distinct neighbourhoods in the central city (e.g. China Town, Little Seoul). In addition, it is strongly represented in newer neighbourhoods in North York, Scarborough and Markham. The South East Asian population exhibits a polar distribution: the poor immigrants are located in the inner city while the wealthy immigrants are located in the suburbs. Immigration for this group has been fairly recent, especially in suburban areas. Specifically, many of the recent immigrants have come from Hong Kong.

Southern Europeans (Yugoslavian, Croatian, Spanish, Greek, Portuguese, Italian)

The Southern Europeans are the second largest single origin ethnic group within Toronto's CMA, e.g. the Italian group is second only to the English. Most of the immigration of the Southern European group occurred after World War II (in the 1950s and 1960s). The Southern Europeans differ from the Western Europeans in that their immigration was mostly sponsored and family related. Thus their initial areas of reception were highly segregated. However, the high levels of segregation have not declined significantly over the years. Southern Europeans have thus tended to resist assimilation. The dominant nationalities within this group are the Italians, Portuguese and Greeks.

Although **Italian** neighbourhoods were previously restricted to the west end of Toronto, they have moved north west over time. Today, much of the Italian population is located in North York, Vaughan, and York. In Vaughan, 49% of the population is of Italian descent (55% of the population is of South European descent). This is mainly because of the strong Italian community in Woodbridge. Notably, Italians have the highest rate of home ownership within Toronto, e.g. over 90% (Ray, 1994).

The **Portuguese** have the second highest rate of segregation within Toronto's CMA; their segregation rate is only slightly less than that of the Jewish. This group has distinct neighbourhoods that are located primarily in the south west areas of Toronto. "The

Portuguese exhibit yet a different pattern of more widely dispersed clusters within the metropolitan area; and the rather compact cluster of high indexes of relative concentration, with very little scatter" (Breton, et al, 1990, p. 97).

The **Greek** population also reveals concentrated clusters within the city. The strongest neighbourhoods are located along the Danforth in the city of East York, where they have 7% of the single origin population. The Greeks, like the other Southern Europeans, reveal moderate to high levels of segregation.

Western Europeans (German, Dutch [Netherlands])

The Western European group has the highest rate of assimilation of any group in Toronto's CMA (Hecht, et al, 1983). Accordingly, this group also has a high rate of intermarriage and English as a first language. Assimilation is facilitated by the fact that people within this group have similar physical characteristics and a greater desire to assimilate. The main immigration period for most Western Europeans was post World War II: immigration was especially high between 1951 and 1955 (Hecht, et al, 1983). Significantly, most of this group immigrated to Canada as skilled labour. "In fact, German immigrants have not been sponsored and that the extended family and kinship system played an insignificant role in the community resulted in a great tendency for the group to disperse geographically across the metropolitan area" (Hecht, et al, 1983, p. 158). Because of their high rate of assimilation, the Dutch and German groups represent only 2.5% of the single origin population. However, when multiple origins are included, they make up 8% (300,960) of the population.

In general, the British, Western Europeans, French and Northern Europeans are considered to be the least segregated ethnic groups. The Eastern Europeans are considered to have intermediate levels of segregation. The East Indians, South East Asians, Southern Europeans, Blacks and Arabs have intermediate to high levels of segregation. The Portuguese and Jewish have the highest levels of segregation.

In this study all indices have been calculated for the following ethnic groups: Aboriginal, Arab (Lebanese), Black, British[ii], Canadian, Chinese, Eastern European, East Indian, English[i], Filipino, French, German, Greek, Italian, Jewish, Northern European, Polish, Portuguese, South East Asian, Southern European, and Western European. Dot distribution maps have also been produced for most of the groups mentioned previously. Other types of maps have been produced for selected groups only. Groups have been chosen because they illustrate the various properties of the maps.

[ii] The British and English groups may not be included if they are used as the comparison group y.

2.3. Scale of Study

The scale of study is Toronto's Census Metropolitan Area (CMA). The CMA is a large urban area, with neighbouring urban and rural areas, that have high degrees of economic and social interaction (1991 Census Dictionary, 1992). In this study, four different levels of aggregation have been used.

a) enumeration area,

b) census tract,

c) federal electoral district, and

d) census subdivision (municipality level).

(a) Enumeration Area:

An enumeration area is the primary and smallest unit of collection for Statistics Canada. Enumeration areas generally follow two criteria: (a) that dwelling populations are between 125 and 375, and (b) that boundaries do not cross a recognized geographic area. However, the boundaries of many enumeration areas change significantly from census to census. For example, 60% of enumeration area boundaries changed from the 1986 to the 1991 Census (1991 Census Dictionary, 1992). In Toronto's CMA there are 5,370 enumeration areas, of which 5,103 have a population that ranges from 40 to 1955, with a mean of 762.545. An enumeration area can represent an area as small as a single apartment building or as large as a rural area encompassing many hamlets. Although the enumeration area is the smallest spatial unit, it is not used as often as census tracts. There are several reasons for this. One reason is that the increased data requirements for enumeration data and boundary files are either too difficult to obtain or too expensive to purchase. Also, although enumeration areas show more detail and variance, they have a greater potential for error, e.g. small populations are more susceptible to outliers. However, this study will measure segregation at the enumeration level in order to assess its usefulness: does the advantage of the detail gained offset the disadvantage of the increased error.

(b) Census Tracts:

Census tracts are the most commonly used spatial units within segregation studies. Their criteria[iii] are that:

(a) "whenever possible, census tract boundaries must follow permanent and easily recognizable physical boundaries;

(b) the population of a census tract must be between 2,500 and 8,000, with a preferred average of 4,000, except for those census tracts in the central business district, in other major commercial and industrial zones, or in peripheral rural or urban areas that may have either a lower or higher population;

(c) when first delineated, or subsequently subdivided, census tracts must be as homogeneous as possible in terms of the economic status and social living conditions of their populations; and

(d) their shape must be as compact as possible" ('1991 Census Dictionary', 1992, pp. 185-86).

However, even with the best intentions, these criteria are not always met. For example, census tracts are generally not homogeneous, e.g. it is nearly impossible to create homogeneous spatial units across

---

[iii] Agreed upon by Statistics Canada and a local committee.

variables. Also, because census tracts are not static entities, they can change from their original state. In Toronto's CMA, there are 812 census tracts of which 809 have a population that ranges from 45 to 14,580, with a mean of 4,809.969. Of 812 census tracts, 704 (~87%) have populations between 2,500 and 8,000.

(c) Federal Electoral Districts

There are 37 federal electoral districts (FEDs) within Toronto's CMA. Some electoral districts extend outside the CMA boundary (e.g. Orangeville, York-Simcoe, Simcoe Centre, Halton-Peel and Ontario). However, in order to keep the study area intact, the portion of these areas which exists within Toronto's CMA has also been included. The population of the electoral districts ranges from 70,150 (Simcoe Centre) to 233,260 (York North). FEDs are important because people who live within a FED are responsible for appointing an official to office. Because of the importance of these boundaries, one should be aware of the possibility of Gerrymandering or the manipulation of boundaries in order to maximize certain populations. "Ethnic spatial concentration can serve what has been termed the 'attack function', a peaceful and legitimate search for, particularly, political representation by a concentration of electoral power" (Fellman, et al, 1990, p. 196). An attack function is considered to be an internal control upon segregation. Including the FEDs is also useful for illustrating the differences between FED and CSD boundaries which are similar in size.

(d) Census Subdivisions

A census subdivision (CSD) can be identified as a municipality or its equivalent, e.g. Indian reserve (1991 Census Dictionary, 1992). In Toronto's CMA there are 27 municipalities or census subdivisions. These range in population from 14,070 (Uxbridge) to 634,770 (Toronto), with a mean of

144,120.93. Also, the range of size of the CSDs is quite broad -- from 15.42 km^2 (Orangeville) to 694.74 km^2 (Caledon). There are several advantages to using CSDs. CSDs are equivalent to cities and are thus recognizable places, e.g. they can be thought of as meaningful units. CSDs are useful for seeing if certain types of maps are suited for high levels of aggregation. Also, CSDs are useful for seeing if ethnic groups are dispersed evenly across political boundaries (also see page 23).

2.4. Issues of Data Quality and Error

Important considerations when using census data are issues of data quality and error. However, it is not known how much an effect error has upon the final outcome of segregation studies using maps and segregation indices. One problem with census data is that increased error occurs with smaller levels of aggregation; this is especially true for enumeration areas. Although census data is considered to be of high quality, one should always be aware of errors and distortions that can occur with this type of data.

Not only can there be errors in representing data, but there can also be error in the data itself. Notably, error implies that there is a true or correct answer. The nature of data determines the expression of error or uncertainty. The nature of error can be broken down into two main types, systematic and random (Eastman et al, 1993). Random error, the most common, is usually associated with measurement error. Systematic error is basically the rate at which error accumulates. Because error is usually thought of as random, statistics can be used to describe and eventually model its effects. In this context, statistics can be thought of as a framework or guideline. Random error itself can be subdivided into positional and attribute error. Although errors can occur in the positional accuracy of census data, these are small compared to attribute errors because their boundaries are well established and fixed. However, positional errors can be quite dramatic when using other data that have corresponding boundaries that do not match. Attribute errors from data collection have been summarized by Census Canada as coverage errors, non-response errors, response errors, processing errors, and sampling errors. Coverage errors are errors that occur when responses are missing; non-response errors are errors that occur when responses cannot be obtained (e.g. because of long periods of absence, irregular working hours, etc.); response errors are errors that occur in the responses themselves (e.g. lying, misreading or misunderstanding questions, etc.); processing errors are errors in the recording of information; sampling errors are errors that result from collecting samples that do not reflect the targeted population. Basically, error measurement or assessment in its simplest form is the classic measure of error or residual being: ERROR = predicted value minus the actual value. Usually this measure is a test to see how well the model performs. There are several standard error descriptors in mapping and GIS. The most common is the RMS or root mean squared error that is used for quantitative data. For qualitative data, error assessment is usually measured as a proportion: those measured correctly against those that are not. However, it is often difficult to obtain such information (it involves going into the field to obtain ground truthing). Although Census Canada checks and adjusts its original figures, census data is rounded off. This rounding only becomes a problem when small population samples are used (e.g. enumeration areas). Also, rounding can become a problem and create a more exaggerated sampling error when a 20% sample is used.

Error propagation will not be performed in this study, however, it is important to consider the possibilities for error when representing ethnic residential segregation.

2.5. Segregation Indices

Residential segregation was measured by using five indices of segregation (recommended by Massey and Denton, 1988) which include the (1) Index of Dissimilarity <Evenness>, (2) Interaction Index <Exposure>, (3) Relative Concentration Index <Concentration>, (4) Absolute Centralization Index <Centralization>, and (5) Spacial Proximity Index <Clustering>. These indices were used to represent the five spatial dimensions of segregation (indicated in inner brackets after each index). Generally, an index value above 0.6 indicates a high level of segregation; between 0.3 and 0.6, an intermediate level of segregation; and below 0.3, a low level of segregation.

Indices were calculated for four different levels of aggregation (enumeration areas [EA], census tracts [CT], census subdivisions [CSD], and federal electoral districts [FED]). In addition, indices dependent upon area were generated from a dasymetric representation (the Relative Concentration and the Absolute Centralization indices were used). The scale that was used was the census metropolitan area (CMA) of Toronto. Results were repeated for the different ethnic groups used.

The following methods were used for calculating the indices.

2.5.1. <EVENNESS>

The Index of Dissimilarity

The dissimilarity index measures the relative concentration of a group across spatial units compared to the proportion of the group for the city as a whole. The behaviour of the index varies between 0 and 1, with one representing absolute segregation and zero representing equal distribution of a group throughout the city. In other words, zero would mean that if the proportion of a group for the entire city was 0.1, then the group would represent 10% of the population in every spatial unit. On the other hand, if a city contained two groups and each had a dissimilarity index value of 1, that would mean that each spatial unit would contain either group A or group B, exclusively. In addition, the dissimilarity index is linked to aggregation because as we approach one person per spatial unit, the index approaches one. Oppositely, if the study area represents one spatial unit, the index values will be zero. Therefore, the smaller the population of each spatial unit, the smaller the dissimilarity index.

In this study, the dissimilarity index was calculated by using the following formula:

$$D = \sum_{i=1}^{n} [t_i|p_i - P|/2TP(1-P)] \qquad [1]$$

where t = total population
p = proportion of minority population
T = total population for the study area
P = proportion of minority population for the study area
i = the spatial unit
n= # of units

The d index was calculated in SPSS. The following is an example of the syntax:

```
COMPUTE d_pol = (tot_pop * ABS((polish/tot_pop) - .0184) / (2 * 3891265 * 0.0184
* (1 - 0.0184)).
```

To finish the calculation, the newly made variables were summed.

2.5.2. <EXPOSURE>

Interaction Index: xP*y

The interaction index shows the probability of group x encountering group y within a given spatial unit. This index varies from 0 to 1, 1 meaning that the probability of a random person from group y encountering a random person from group x would be 100%. Therefore, a xP*y value of one would be theoretically impossible unless group x and group y were the same. The minimum, zero, would mean that the probability of group y encountering group x was zero percent within a given spatial unit. The two extremes are highly unlikely. A more realistic value would read as follows. A value of 0.4 would mean that there was a 40% chance of a member of group x running into a member of group y within a given spatial unit. The interaction index is very dependent upon group size. Also, the index is asymmetrical, meaning that if groups x and y are reversed, the xP*y values will be dissimilar, unless each group is identical in size and distribution. The interaction index has been criticized because it is asymmetrical and strongly affected by group size. However, since we are trying to measure isolation, the size of a group becomes very important.

The interaction index was calculated by using the following formula :

$$xP^{*}y = \sum_{i=1}^{n} [x/X][y/t_i] \qquad [2]$$

where x = population of group x
y = population of group y (comparison group)
X = population of group x for entire study area
t = total population
i = the spatial unit
n= # of units

The xP*y index was calculated in SPSS as well. The following is an example of the syntax:

```
COMPUTE bxy_germ = (german / 66875) * (british / tot_pop).
```

The newly created variables were also summed. The English and the British variables were used as the comparison group y. The two were included to illustrate the affect of group size.

The isolation index (xP*x) was also calculated for the exposure dimension. The isolation index is basically identical to the interaction index, except group y is replaced by group x. Therefore, if group x has a value of 0.8, group x would have an 80% chance of running into another person from the same ethnic group, within their own spatial unit.

The isolation index was calculated by using the following formula :

$$xP^{*}x = \sum_{i=1}^{n} [x_i/X][x_i/t_i] \qquad [3]$$

where x = population of group x
X = population of group x for entire study area
t = total population
i = the spatial unit
n= # of units

2.5.3. <CONCENTRATION>

Relative Concentration Index: RCO

The relative concentration index (RCO) measures concentration by measuring the sum of the proportion of group x, times the area of each spatial unit, divided by the sum of the proportion of group y, times the area of each spatial unit, minus one. The nominator measures the average amount of area occupied by group x, divided by the average amount of area occupied by group y. The minus one ensures that the value does not go above one. The denominator measures the minimum area that group x could occupy over the maximum area that group y could occupy. The denominator is also subtracted by 1. "The relative concentration index measures the share of urban space occupied by group x compared to group y" (Massey and Denton, 1988, p. 291). The RCO index varies from -1 to +1, with zero suggesting that "the two groups are equally concentrated in urban space. A score of -1 means that y's concentration exceeds x's to the maximum extent possible, and a score of +1 means the converse" (Massey and Denton, 1988, p. 291).

The calculation of the RCO index involved several steps:

1. finding the area for each spatial unit;
2. finding the cumulative sum for the population of each group, sorted in ascending order by area,
3. calculating the formula.

1. The calculation of area for each spatial unit was done in Mapinfo.

2. To find the cumulative sum for the population of each spatial unit by area, a Qbasic program was used (see Appendix). The Qbasic program calculated a cumulative sum and gave a case number, but did not sort variables. Therefore, the population variable was first sorted by area and then exported. It

was then run through the Qbasic program and brought back into SPSS.

3. The index was calculated by using the following formula:

$$RCO=([\sum_{i=1}^{n}(x_i a_i/X)/\sum_{i=1}^{n}(y_i a_i/Y)-1]/[\sum_{i=1}^{n_1}(t_i a_i/T_1)/\sum_{i=n_2}^{n}(t_i a_i/T_2)-1]) \quad [4]$$

where a_i = area of unit i
Y = population of group y for entire study area (comparison group)
y = population of group y (comparison group)
t = total population
T = total population for the study area
x = population of group x
X = population of group x for entire study area
i = the spatial unit
n= # of units
T_1 = total population from tracts 1 to n_1, "where n_1 is the rank of the tract where the cumulative total population = the minority population for the whole city" (Massey and Denton, 1988, p. 290).
T_2 = total population from tracts n_2 to n, "where n_2 represents the tract where the cumulative total population = the minority population for the whole city" (Massey and Denton, 1988, p. 290).

This equation can be simplified into the following:

$$RCO = ((A / B) - 1) / ((C / D) - 1) \quad [5]$$

Where A, B, C and D are illustrated by the following SPSS syntaxes:

```
COMPUTE a_bla = (black * area) / 124560 .
COMPUTE b_eng = (english * area) / 516060 .
IF (tot_cum <= T1 ) c_bla = (tot_pop * area) / T1.
COMPUTE tot_cum1 = 3891265 - tot_cum
IF (tot_cum1 > 124560) d_bla = (tot_pop * area) / T2 .
```

Each variable was then summed and brought into Equation 5 to be substituted. The English and British variables were used as group y.

2.5.4. <CENTRALIZATION>

Absolute Centralization Index: ACE

The ACE index measures centralization. Centralization can be defined as "the degree to which a group is spatially located near the centre" (Massey and Denton, 1988, p. 293). The ACE index ranges from +1 to -1, with zero representing uniform distribution across the city, +1 representing a tendency for group x to live close to the Cental Business District (CBD), and -1 representing a tendency for group x to live in outlying areas. Therefore, if group x had an index value of 0.25, that would mean that 25% of group x would have to move further from the CBD in order to achieve uniform distribution around the CBD. Another example would be that if group x had a value of -0.1, this would mean that 10% of group x would have to move closer to the CBD in order to achieve uniform distribution around the CBD.

However, the ACE index does not tell how close a group is to the CBD. In addition, this index does not tell us anything about where a group has come from or should go to in order to achieve uniform distribution.

In this study, the calculation of the ACE index involved the three following steps:

1. finding the distance from each area unit centroid to the CBD and the area of each areal unit;
1. finding the cumulative sum for each ethnic variable and area;
2. calculating the index.

1. The distance from each areal unit to the CBD was calculated in Mapinfo. The start coordinate (-79.38, 43.655) represents the CBD at the intersection of Bay and King. The geographic projection (latitude, longitude) was thought to be sufficient because only the relative distance was needed. The results were then exported into a *.dbf file and brought into SPSS.

2. The spatial units were then ranked by distance from the CBD in ascending order. The cumulative sum was accomplished by using the same Qbasic program that was used for the RCO index. The ethnic and area variables were isolated and exported into text files to be run through the Qbasic program. The output files were then brought back into SPSS.

The following formula was used to calculate the ACE index:

$$ACE = \left(\sum_{i=1}^{n} X_{i-1} A_i\right) - \left(\sum_{i=1}^{n} X_i A_{i-1}\right) \quad [6]$$

where A_i = cumulative proportion of area through unit i
[X_i = cumulative proportion of the population of group x through unit i]
i = the spatial unit
n= # of units

As with the other indices, the calculations were done in SPSS. The following is an example of the syntax used:

```
compute ace_seur = (lag(seur_cum) * area_cum) - (lag(area_cum) * seur_cum).
```

This was repeated for all ethnic variables. All variables were then summed to complete the equation.

2.5.5. <CLUSTERING>

Spatial Proximity Index

The spatial proximity index compares the relative degree of spatial clustering of group x to group y. This index addresses the problem of what is known as the 'checkerboard problem' (White, 1983). Large, continuous ethnic enclaves (or ghettos) are generally thought to be worse than small scattered ethnic enclaves. The spatial proximity index is very similar to autocorrelation measures which measure spatial concentration by a contiguity matrix. While the spatial proximity index also measures concentration through a contiguity matrix, it also takes into account the concentration of group y.

The "Spatial Proximity Index is simply the average of intragroup proximities, P_{XX}/P_{TT} and P_{YY}/P_{TT}, weighted by the fraction of each group in the population" (Massey and Denton, 1988, p. 295). This index ranges from 0 to 1. A value of zero would mean that there was a perfect checkerboard distribution, while a value of one would mean that all the people in group x would be located next to one another. The connectivity approximation c_{ij} becomes very small when the measurement goes above one. When the distance between spatial units i and j is zero (the same spatial unit), c_{ij} equals 1; however, when spatial units i and j move farther apart, c_{ij} equals zero. Therefore, the choice of unit measurement is important, especially for comparative purposes.

In this study, the calculation of the spatial proximity index involved the following two steps:

1. creating a distance matrix
2. calculating the index.

1. The first step was done in ArcInfo. A district matrix was produced by the POINTDISTANCE command in ArcInfo. This command creates a file of distance values from each polygon centroid to every other polygon centroid in the coverage. If there are a large number of polygons in the coverage, a huge file would result. For example, an EA coverage of 5,370 enumeration areas would result in 28,836,900 records. Obviously, this would be unmanageable. Therefore, in this study, a search radius of 1km was used. Since $c_{ij} = \exp(-d_{ij})$, c_{ij} goes down drastically after one. This produced a table with ~152,000 records -- still quite large.

2. Also, the calculation of the spatial proximity index was done in SPSS by using the following formula:

$$SP=(XP_{xx}+YP_{yy})/TP_{tt} \tag{7}$$

where P_{xx} = the average proximities between x members
P_{yy} = the average proximities between y members
P_{tt} = the average proximities between t members
X = population of group x for the entire study area
Y = population of group y for the entire study area
T = total population for the entire study area

and P_{xx} was measured by

$$P_{XX}=\sum_{i=1}^{n}\sum_{j=1}^{n}x_i x_j c_{ij}/X^2 \tag{8}$$

where c_{ij} = one element in a contiguity matrix -- can be estimated by $c_{ij} = \exp(-d_{ij})$,
where d = distance
X = population of group x for the entire study area
x = population of group x in spatial unit i or j
j = neighbouring spatial unit
i = the spatial unit
n= # of units

P_{YY} and P_{TT} are calculated the same way, except group y and the total population values are substituted for x and X. The following are examples of SPSS syntaxes used in the calculation of the spatial proximity index.

```
COMPUTE c = EXP(-distkm) .
COMPUTE pxx_jew = (i_jewis * j_jewis * c) /  113260 ** 2 .
COMPUTE pyy_eng = (i_english * j_english * c) / 516060 ** 2 .
COMPUTE p_tt = (i_totpop * j_totpop * c) / 3891265 ** 2 .
```

After the P_{XX} values were calculated for the different ethnic groups, all variables were summed. The values that resulted were then brought into Equation 6 and substituted.

2.6. Cartographic Representation

All cartographic representation was performed with MapInfo, Arcview or ArcInfo software. The ethnic spatial distribution was mapped as a proportion of the total population of each spatial unit. The boundary files were obtained from Wilfrid Laurier University (census tracts and above) and the enumeration boundary files were obtained from Compusearch Micromarketing Data Systems in Toronto.

The dasymetric technique was achieved by delineating the residential land within Toronto's CMA. Most of the residential delineation was obtained from Compusearch. However, areas northwest of the city were added and areas within the GTA were corrected or enhanced where needed. For example, small enumeration areas that were missing, as well as populated rural areas, were added. Also, considerable editing was done to both the land use and census boundaries so that their shared boundaries would match. This was essential to perform Boolean AND overlays. The new residential polygons were digitized. Head up digitizing was also done by snapping to street layers. Reference material included Smart maps, municipal maps, Rand-MacNally maps, SNF street files, topographic maps, etc. The scale of the maps ranged from 1:500 to 1:50000 in less populated areas where there was little large scale coverage.

After the residential coverage was produced, the different levels of aggregation were intersected in ArcInfo with the UNION command (Boolean AND overlay). This produced a file that included census boundaries and residential designations, and preserved all attributes from each of the original coverages. Notably, many spurious polygons were created. This was corrected by dissolving the polygons according to the EA id. The non-residential areas were then either deleted or merged together. Since some enumeration areas encompassed more than one polygon, the polygon coverages were converted or aggregated into regions.

All info tables with new area measurement were exported as *.dbf files and brought back into SPSS. The indices that had an area variable were then recalculated.

All the maps displayed either absolute values or ratios. There were three exceptions:

- nominal areal map
- location quotient map
- cartogram.

The following methodology was used for creating these maps.

2.6.1. Nominal Areal Map

A dominant ethnic group map was produced by classifying spatial units into the dominant ethnic groups. The first pass was classified by allocating spatial units to an ethnic group if it represented over 30% of the population. Spatial units were also classified into a dominant ethnic group if they had an abnormally high representation in an area. Spatial units without a dominant group were labeled as 'mixed'. Classifying spatial units into dominant ethnic groups clarified the cartographic message by reducing the amount of information within the map.

2.6.2. Location Quotient Map

Like other indices mentioned in this chapter, the location quotient is a descriptive spatial statistic. The location quotient compares "some quality of an area with the specified norm" (Blakemore, in Johnson, et al, p. 299). The location quotient values are summarized in the following chart:

$LQ > 1$	more share of activity
$LQ = 1$	equal share of activity
$LQ < 1$	less share of activity

Basically, a value over one means that the phenomenon is over-represented in an area, while a value of less than one means that the phenomenon is under-represented in an area.

The location quotient (LQ) was calculated by using the following formula :

$$LQ = (x/X)/(t/T) \qquad [8]$$

where x = population of group x in spatial unit i
t = total population in spatial unit i
X = population of group x for entire study area
T = total population
i = the spatial unit
n= # of units

Location quotient values were then mapped in a standard mapping package.

2.6.3. Cartogram

A population cartogram was produced as an alternative to using standard arbitrary census boundaries. The cartogram produced in this study was a very simple cartogram. The CSDs were represented as circles that were proportional to the total population of the CSD. Therefore, CSDs with the highest populations were visually dominant. These circles were then shaded like a regular proportional map. However, because of the limitations of the software, the proportional circles were exported as a bitmap to another application, where appropriate colours could be added.

CHAPTER 3

SEGREGATION INDICES

3.1. Segregation Dimensions and Indices

Since indices are a common way of acknowledging and confirming segregation, their sensitivity is important to the way that we perceive and identify segregation. This chapter will look at the effect of aggregation upon indices (recommended by Massey and Denton, 1988) that represent the five dimensions of segregation. These indices will be calculated on four standard levels of aggregation (EA, CT, FED and CSD). Discussion will be broken down into three main parts:

a. **Index Values:** a comparison of the index values of the different ethnic groups at four standard levels of aggregation;

b. **Significance Test:** a testing of the index values to see if they change significantly at different levels of aggregation;

c. **Segregation Ranks:** a comparison of the segregation ranks at different levels of aggregation;

In addition, the concentration and centralization dimensions will include an additional section. Because the ACE and RCO indices have area in their equations, they will be recalculated with residential area only to see how sensitive these indices are to area. Therefore, this study will also include:

d. **Residential Calculations:** a comparison of the original index values with index values that use residential area only.

3.1.1. Evenness

Evenness refers to how evenly a group is distributed over space. The evenness dimension is the dimension that is most often associated with segregation. In this study, the dissimilarity index has been used to measure the evenness dimension. Notably, the dissimilarity index has been used as the standard for measuring segregation for over 40 years. In Table 3.11 the dissimilarity index was calculated on single-origin ethnic groups within Toronto's CMA at four levels of aggregation.

3.1.1.1. Index Values

Table 3.11: Dissimilarity index for Toronto's Census Metropolitan area (1991) at four levels of aggregation

Ethnic Group	ea (5370)	ct (812)	fed (37)	csd (27)
Aboriginal	0.912	0.595	0.205	0.170
Arab (Lebanese)	0.901	0.623	0.335	0.300
Black	0.571	0.425	0.283	0.214
British	0.415	0.242	0.166	0.130
Canadian	0.334	0.235	0.173	0.152
Chinese	0.595	0.509	0.401	0.318
Eastern European	0.473	0.365	0.299	0.250
East Indian	0.562	0.430	0.302	0.233
English	0.295	0.229	0.153	0.127
Filipino	0.620	0.421	0.237	0.179
French	0.452	0.240	0.127	0.113
German	0.416	0.239	0.120	0.089
Greek	0.618	0.442	0.302	0.242
Italian	0.568	0.495	0.403	0.312
Jewish	0.811	0.775	0.669	0.548
Northern European	0.786	0.423	0.170	0.120
Polish	0.592	0.453	0.369	0.279
Portuguese	0.656	0.577	0.474	0.414
South East Asian	0.435	0.414	0.313	0.252
Southern European	0.535	0.526	0.525	0.525
Western European	0.395	0.252	0.153	0.151

The dissimilarity index ranges between 0 and 1, with one representing absolute segregation and zero representing equal distribution of a group throughout the study area. In Table 3.11, the Jewish group had a value of 0.811 at the EA level, which meant that 81.1% of the Jewish group would have to move to spatial units where they were under represented in order to achieve uniform distribution. According to Denton (1994), Breton et al. (1990), etc., a value of 0.7 and above is considered to be segregated. As expected, the d-index values for most of the groups increased as the number of units increased. This is reasonable since variance decreases as information is lost. On average, the highest values came from the Jewish and the Portuguese, while the lowest values came from the Western Europeans and the English. This is consistent with previous findings. However, at the EA level, the values of the smallest groups [e.g. the Northern Europeans, Aboriginals, and Arabs (Lebanese)] exceeded the values of the Jewish and Portuguese.

3.1.1.2. Significance Test

A Friedman Two-way Anova test was performed to see if the aggregation levels were significantly different. The Friedman test is a non-parametric rank-order test that is used for related groups. The Friedman test was used instead of a parametric test because: (a) the samples were not normally distributed; (b) the samples had less than 30 cases; and (c) the samples were not independent. Although the Friedman test is not as powerful as a parametric test, its assumptions are more relaxed. A comparable parametric test is the t-test. "A parametric test compares the means of the two groups; its null hypothesis is that the two populations have the same mean. The equivalent to the mean in a rank-order test is the middle rank, which is the median of the non-ranked scores.... Thus a rank-order test compares the medians of two groups, its null hypothesis is that the two populations have the same median" (Aron and Aron, 1994, p. 259). A rank order test does this by ranking all the values and then calculating the mean rank score for each sample. The mean ranks are then compared along a chi square distribution.

Four samples were used for the Friedman test:

Sample one. Enumeration Area (EA)

Sample two: Census Tract (CT)

Sample three: Federal Electoral District (FED)

Sample four: Census Subdivision (CSD)

The null hypothesis of this test (H_0) was that all the dissimilarity index values stayed the same when the data was aggregated. Therefore, H_1 was that the dissimilarity index values changed significantly when the data was aggregated.

Table 3.12: Friedman Two-Way Anova Test comparing Dissimilarity Index values.

- - - - - Friedman Two-Way Anova

Mean Rank	Variable
1.02	CSD
3.00	CT
4.00	EA
1.98	FED

Cases	Chi-Square	D.F.	Significance
21	62.713	3	.0000

Since the Friedman test returned a low significance value, H_o was rejected. Therefore, the dissimilarity index values changed significantly when the data was aggregated.

3.1.1.3. Segregation Ranks

"In most situations, segregation measures are used to compare levels of segregation in different regions. Thus, one may argue that if the scale effect imposes similar impacts on every region, there is no need to worry about the magnitude of the effect, and scale effect will not change the relative levels of segregation" (Wong, 1997, p. 131).

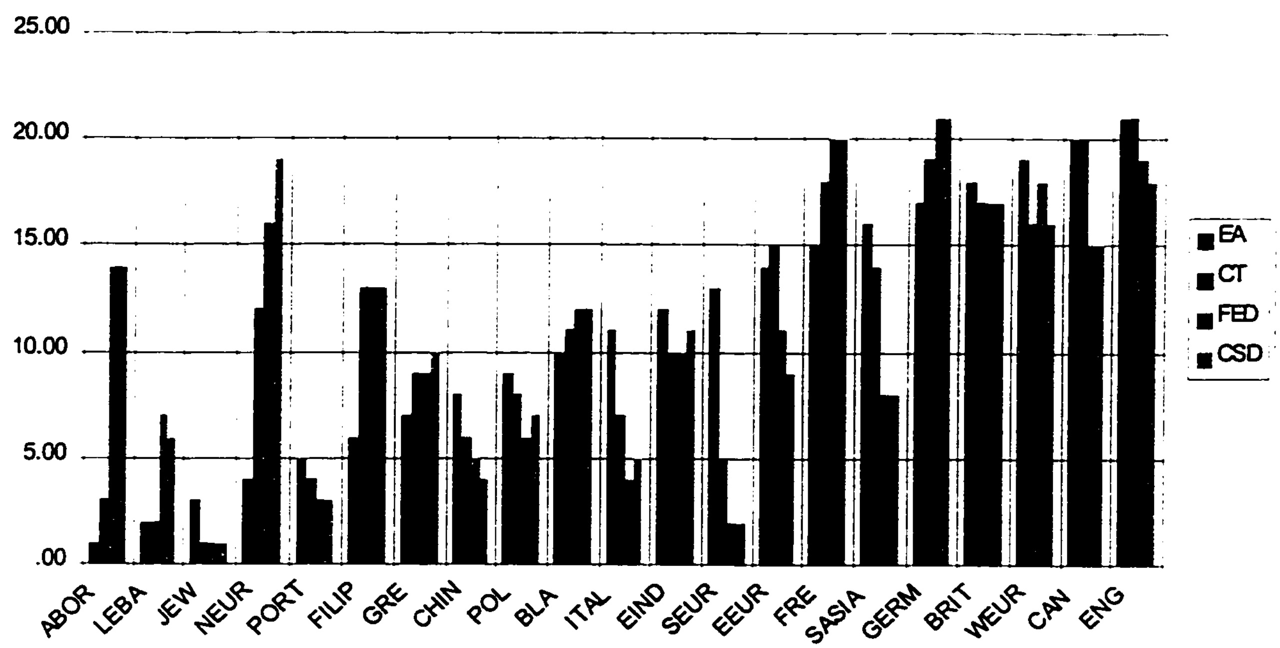


Figure 3.1: Dissimilarity index ranks at four standard aggregation levels.

To further explore the difference between the aggregation levels, the rankings of the ethnic groups were compared individually. Figure 3.1 shows the dissimilarity index rankings for different ethnic groups at four levels of aggregation. Accordingly, the segregation ranks stayed relatively the same for most of the ethnic groups. The following groups were exceptions: (i) the Northern Europeans, Aboriginals, and Arabs (Lebanese), and (ii) the Southern Europeans.

(i) The dissimilarity index values of the three smallest groups (the Northern Europeans, Aboriginals and Arabs [Lebanese]) were interesting because they changed more dramatically than the values of the other ethnic groups. For example, the Northern European group went from being the second most highly segregated group (.79 [EA]), to being one of the least segregated groups (.12 [CSD]). The Northern European group was particularly interesting because the dissimilarity index values at the EA level were not consistent with previous findings. Previous findings have shown that the Scandinavian group is a highly integrated group (similar to the Western Europeans).

The most likely explanation for the dramatic change in rank of the smaller groups was the effect of group size upon the index. Since there are less people per unit at the EA level, a slight increase in population could increase proportion and variation within the spatial unit. When groups have a small proportion of the total population, they are more affected by slight population changes within the spatial unit. Theoretically, as the number of spatial units becomes smaller, the dissimilarity index moves closer to one (e.g. each spatial unit represents one person). Also, sampling error could have been a consideration. The increased sampling error that accompanies small sample size makes an index more unstable.

Another explanation could have been that the dissimilarity index was correct. This would mean that according to the evenness dimension, small ethnic groups were highly segregated within small census units -- such as apartment buildings or neighbourhood blocks. However, as the size of spatial units increased and the index values decreased, small pockets were cancelled out by areal averaging. These pockets were too small and too scattered to constitute ethnic enclaves.

(ii) The other group that behaved differently was the Southern European group. The measured level of segregation of this group (.53) remained constant across different levels of aggregation. However, the Southern European group went from a rank of 11 out of 19 at the EA level, to a rank of 2 out of 19 at the CSD level. This pattern is opposite the pattern of the smaller groups. One reason for the constant rate of segregation of this group is that it is strongly represented in large areas of the city that correspond to census tract, federal, and municipal boundaries. This is consistent with earlier findings. However, if we look at the individual ethnic groups that make up the Southern European group, we can see that the values of the dissimilarity index go down with increased levels of aggregation. This decrease, however, is not as extreme as with other groups. In addition, because of the large population of this group (second only to the British), their dissimilarity index did not have the same instability problems as the dissimilarity index of the smaller groups.

The rate at which the dissimilarity index decreases can therefore highlight various properties of distribution. Extreme differences between index ranks at different levels of aggregation are the result of areal averaging that cancels out groups that are segregated in small, scattered units throughout the city. Conversely, the less the difference between aggregation levels, the more the likelihood that there are large established ethnic enclaves. The difference between dissimilarity indices, at different levels of aggregation, can tell how scattered a group is or whether a group is equally represented in municipalities or electoral ridings. However, the dissimilarity index, like other indices, fails to show location. Maps therefore complement findings because they show both area and location (see Chapter 4).

3.1.2. Exposure

"The Exposure dimension, through P-type measures, seeks to describe the relative isolation of groups in terms of probability models of interaction among themselves and with others" (Leiberson and Carter, 1988, p. 297). The exposure dimension is more concerned with the experience of segregation than with the mechanical aspects of segregation. Segregation is considered more extreme when it is accompanied by isolation. Massey and Denton (1988) have recommended the interaction index to represent the exposure dimension. The interaction index measures exposure as the probability of group x encountering group y within a given spatial unit. This index varies from 0 to 1, with 1 suggesting that the probability of a random person from group x encountering a random person from group y is 100%. Another related p-type index is the isolation index. The isolation index "measures the extent to which minority members are exposed to one other, rather than to majority members, and it is computed as the minority-weighted average of each unit's minority proportion" (Massey and Denton, 1988, p. 288). The isolation index (which also varies between 0 and 1) is read the same as the interaction index except group y is replaced by group x. According to Massey and Denton (1988), segregation is present with an interaction index value of 0.6 and above. However, the higher the interaction index value, the less the isolation and the more exposure group x has to the core group y. Consequently, high interaction index values describe exposure and not isolation. Conversely, high isolation index values describe isolation. For this reason, both the interaction and the isolation indices have been calculated.

3.1.2.1. Index Values

Table 3.21: The Interaction index ($xP^{*}y$) for Toronto's CMA (1991) at four standard levels of aggregation. English is the comparison group y, unless stated otherwise.

GROUP X	EA xP*y	CT xP*y	FED xP*y	CSD xP*y	EA_brit xP*y	CT_brit xP*y	FED_brit xP*y	CSD_brit xP*y
Aboriginal	0.141	0.138	0.135	0.132	0.208	0.2	0.196	0.191
Arab	0.119	0.124	0.119	0.122	0.169	0.181	0.174	0.178
Black	0.102	0.107	0.119	0.123	0.142	0.153	0.172	0.18
Canadian	0.159	0.158	0.148	0.148	0.228	0.226	0.213	0.211
Chinese	0.086	0.094	0.114	0.12	0.127	0.139	0.167	0.175
East Indian	0.091	0.103	0.124	0.127	0.186	0.189	0.187	0.185
Eastern European	0.127	0.13	0.129	0.127	0.13	0.15	0.18	0.185
Filipino	0.1	0.109	0.126	0.125	0.146	0.159	0.184	0.182
French	0.15	0.146	0.14	0.138	0.218	0.212	0.202	0.199
German	0.157	0.154	0.144	0.142	0.227	0.222	0.207	0.204
Greek	0.111	0.116	0.124	0.124	0.136	0.169	0.181	0.182
Italian	0.09	0.096	0.111	0.111	0.127	0.135	0.158	0.159
Jewish	0.072	0.078	0.109	0.095	0.108	0.115	0.158	0.137
Northern European	0.159	0.152	0.14	0.139	0.229	0.221	0.202	0.199
Polish	0.121	0.125	0.126	0.126	0.176	0.183	0.184	0.183
Portuguese	0.088	0.093	0.104	0.124	0.124	0.131	0.149	0.18
South E. Asian	0.046	0.01	0.117	0.121	0.052	0.146	0.17	0.176
Southern European	0.094	0.099	0.112	0.117	0.133	0.141	0.16	0.168
Western European	0.161	0.159	0.149	0.148	0.232	0.227	0.213	0.211

Table 3.21 shows the interaction index ($xP^{*}y$) at four levels of aggregation. With the interaction index, the use of both the British and the English as group y, illustrates the effect of changing the population of group y. In all cases, the probability was higher that each group would encounter the larger British group than the smaller English group within their own spatial unit. The sensitivity to group size has been an argument against the use of P-type indices. However, there have also been arguments that support this aspect of P-type indices (see Lieberson & Carter, 1988). An important argument has been that the size of a group affects the perceived effect of segregation and therefore influences the probability of one group encountering another within their own spatial unit. Another property of P-type

indices that is related to population size, is that the indices are assymetrical. In other words, $xP^{*}y <> yP^{*}x$, unless groups x and group y have the same population. For example, if group x had 10% of the population and group y had 90% of the population, each would have a different experience. Therefore, different index values reflect different experiences.

As indicated in Table 3.21, all the values of the interaction index were relatively low. For example, people of single German origin had a 22.7% chance (EA) of running into someone of single British origin within their own spatial unit. The Chinese, South East Asians, and Jewish had the lowest probability of encountering the English and/or the British within their own spatial units. Accordingly, none of the ethnic groups showed a high probability of running into the English or British within their own spatial units. The small interaction index values[i] experienced in Toronto can be attributed to the relative size of the ethnic groups. Toronto is a large and diverse multicultural society where each ethnic group is but a small part of broad and diverse whole. Thus the variety and consequent diminished size of each ethnic group impacted upon the interaction index. Also, the use of single-origin ethnics instead of racial groups had an impact because it further reduced group size.

---

[i] In contrast, the following are the average interaction index values for American cities: American Indian - 0.874, Asian - 0.690, Black - 0.376, Hispanic - 0.474, Non-Hispanic Black - 0.377 (From ... http://www.census.gov/pub/hhes/www/housing/resseg/sumtabs.html.).

Table 3.22: The Isolation index (xP*x) for Toronto's CMA (1991) at four standard levels of aggregation. English is the comparison group y

GROUP X	EA xP*x	CT xP*x	FED xP*x	CSD xP*x
Aboriginal	0.027	0.006	0.002	0.002
Arab	0.049	0.012	0.004	0.003
Black	0.126	0.074	0.049	0.04
Canadian	0.111	0.088	0.079	0.078
Chinese	0.236	0.186	0.118	0.088
East Indian	0.109	0.077	0.064	0.047
Eastern European	0.129	0.085	0.053	0.046
Filipino	0.078	0.04	0.023	0.02
French	0.034	0.018	0.014	0.014
German	0.037	0.023	0.019	0.019
Greek	0.079	0.05	0.031	0.024
Italian	0.315	0.262	0.151	0.143
Jewish	0.34	0.286	0.112	0.083
Northern European	0.025	0.007	0.004	0.004
Polish	0.097	0.056	0.042	0.026
Portuguese	0.223	0.181	0.124	0.058
South East Asian	0.051	0.197	0.139	0.117
Southern European	0.35	0.303	0.22	0.182
Western European	0.056	0.036	0.03	0.03

The values of the isolation index were more varied than the values of the interaction index (see Table 3.22). This was because the isolation index measured the probability of running into someone from the same group (group x), while the interaction index measured the probability of running into someone from group y. As expected, the groups with the lowest interaction index values had the highest isolation index values. The highest isolation index values came from the Jewish (0.34), Italians (0.314), Chinese (0.236), and Portuguese (0.223); while the lowest values came from the South East Asians (0.051), Arabs (0.049), Germans (0.037), French (0.034), Aboriginals (0.027), and Northern Europeans (0.025).[ii]

[ii] For comparison, the following are the average isolation index values for American cities: American Indian - 0.126, Asian - 0.310, Black - 0.624, Hispanic - 0.526, Non-Hispanic Black - 0.623. (From ... http://www.census.gov/pub/hhes/www/housing/resseg/sumtabs.html.).

With both the interaction and isolation indices, the variance decreased as the number of spatial units decreased. This was because the interaction index moved towards the population proportion of group y and the isolation index moved towards the population proportion of group x. For example, if the study area was one spatial unit, the interaction index would simplify into y/t, where y equals the population of group y and t equals the total population; and the isolation index would simplify into x/t, where x equals the population of group x and t equals the total population. To illustrate this further, the interaction index was looked at in greater detail. The interaction index moved closer to a value of 0.133 with English as the comparison group y, and 0.18 with British as the comparison group y. The values of 0.133 and 0.18 reflected the population percentage of group y. If the study area had been divided into one spatial unit, all ethnic groups would have had an interaction index value of 0.133 (with the English as group y). Therefore, when looking at the interaction index, groups with values above 0.133 had a higher probability of running into the English within their own spatial units than within the whole CMA; oppositely, groups with index values below 0.133 had a lower probability of running into the English within their own spatial units than within the whole CMA. Another way of looking at the interaction index would be to write the $_xP^*_y$ index as follows: $_xP^*_y = (E[x/X][y/t]) - [y/t]$. The addition of y/t to the end would mean that as the number of spatial units decreased, the index would approach zero. Using this new equation, a group with a normal interaction index value below 0.133 would have a negative number, suggesting that this group was less likely to run into the English within their own spatial unit than within the whole CMA. Notably, this rewriting of the index has been done for illustrative purposes only. There has been no recommendation to alter the index because, if altered, the index would no longer vary between 0 and 1. It was felt that an index range of 0 to 1 was essential for conveying probability.

3.1.2.2. Significance Test

The Friedman test was again used to assess whether the index values from the different levels of aggregation were significantly different. In Table 3.23, the Friedman test showed that the null hypothesis was not rejected because the samples were not significantly different (e.g. 0.3100).

Table 3.23: Friedman Two-Way Anova Test comparing xP*y values.

- - - - - Friedman Two-Way Anova

Mean Rank	Variable
2.05	EA
2.47	CT
2.76	FED
2.71	CSD

Cases	Chi-Square	D.F.	Significance
19	3.5842	3	.3100

The samples were not significantly different because the Friedman test compared the median or middle index value of all the samples. As mentioned previously, although the index values all disperse around 0.133 (with the English as group y), there is less variance around 0.133 when the data is aggregated. However, we are comparing median scores, not variance. Even though the variance might change radically, the median or middle scores might stay the same. For example, if there were two samples, Group A (e.g. 1,2,3,4,5) and Group B (e.g. 3,3,3,3,3), both would have the same median value but both would be different samples. The interaction index behaves in much the same way. The more spatial units there are, the more the sample looks like Group A [e.g. 1,2,3,4,5], and the less spatial units there are, the more the sample looks like Group B (e.g. 3,3,3,3,3). For other indices, the values usually decrease (e.g. the dissimilarity index) or increase (e.g. the SP index) when the number of spatial units is changed. This causes the median to change considerably and therefore return a low significance value.

3.1.2.3. Segregation Ranks

Another and perhaps more meaningful evaluation was to compare the xP*y ranks individually between the four levels of aggregation. Table 3.24 below shows the interaction index ranks for the various ethnic groups at four levels of aggregation.

Table 3.24: xP*y Segregation Ranking (English group y)

Rank	ea	ct	fed	csd
1	South E. Asian	South E. Asian	Portuguese	Jewish
2	Jewish	Jewish	Jewish	Italian
3	Chinese	Portuguese	Italian	Southern European
4	Portuguese	Chinese	Southern European	Chinese
5	Italian	Italian	Chinese	South E. Asian
6	East Indian	Southern European	South E. Asian	Arab
7	Southern European	East Indian	Black	Black
8	Filipino	Black	Arab	Portuguese
9	Black	Filipino	East Indian	Greek
10	Greek	Greek	Greek	Filipino
11	Arab	Arab	Filipino	Polish
12	Polish	Polish	Polish	East Indian
13	Eastern European	Eastern European	Eastern European	Eastern European
14	Aboriginal	Aboriginal	Aboriginal	Aboriginal
15	French	French	French	French
16	German	Northern European	Northern European	Northern European
17	Canadian	German	German	German
18	Northern European	Canadian	Canadian	Canadian
19	Western European	Western European	Western European	Western European

According to Table 3.24, the xP*y segregation rank for most of the ethnic groups changed very little between the different levels of aggregation. Significantly, the xP*y rank of most of the groups stayed in the same top, middle or bottom third. Therefore, in terms of relative standing, there was little to be found by looking at the different levels of aggregation for most of the groups. However, the segregation rank of some of the groups changed considerably: (i) the Portuguese group went from a rank of 1 (FED) to 8 (CSD); (ii) the East Indian group went from a rank of 4 (EA) to 12 (CSD); and (iii) the Arab group went from a rank of 11 (EA) to 6 (CSD). The ranking change of the Portuguese and the East Indian groups suggested that, on a enumeration or neighbourhood (CT) level, these groups had little exposure to people of English descent within their own spatial units. With the Arab group, the probability of running into someone of English descent remained nearly the same across different levels of aggregation. Although the absolute probability of the Arab group was stable, their relative position increased when the absolute probability of the other groups decreased.

The groups with the lowest rankings (highest exposure) were the Northern Europeans, Canadians, Western Europeans, Germans, and French. These 'core groups' (Hecht et al., 1983) had the highest probability of running into someone of English descent within their own spatial units. The groups with intermediate levels of exposure were the Aboriginals, Eastern Europeans, Filipinos, Arabs, Greeks, and East Indians. The groups with the lowest probability of running into the English were the Jewish, South East Asians, Chinese, and Portuguese.

3.1.3. Concentration

Concentration refers to the population density of a group. Concentration is considered important because "segregation has traditionally restricted minorities to small, densely packed communities" (Denton, 1994, p. 54). High density neighbourhoods are often associated with a low standard of living. To measure the concentration dimension, the relative concentration index (RCO) was used. The RCO index measured concentration by measuring the sum of the proportion of group x, times the area of each spatial unit, divided by the sum of the proportion of group y, times the area of each spatial unit, minus one. In essense, "the relative concentration index measures the share of urban space occupied by group x compared to group y" (Massey and Denton, 1988, p. 291). The RCO index varies from -1 to +1, with 0 suggesting that "the two groups are equally concentrated in urban space. A score of -1.0 means that y's concentration exceeds x's to the maximum extent possible, while a score of [positive] 1 (+1) means the converse" (Massey and Denton, 1988, p. 291). Table 3.31 below shows the RCO values of various ethnic groups at different levels of aggregation.

3.1.3.1. Index Values

Table 3.31: RCO Index for Toronto CMA (1991) at Four Standard Levels of Aggregation

Ethnic Group	ea (5370)	ct (812)	fed (37)	csd (27)
Aboriginal	0.467	0.530	0.340	0.140
Arab (Lebanese)	0.759	0.693	0.610	-0.010
Black	0.810	0.716	0.610	0.050
Canadian	0.031	0.032	-0.060	-0.040
Chinese	0.769	0.733	0.520	0.110
Eastern European	0.549	0.496	0.530	0.100
East Indian	0.718	0.587	0.460	-0.060
English	0.000	0.000	0.000	0.000
Filipino	0.697	0.626	0.500	0.000
French	0.472	0.433	0.260	0.050
German	-0.200	-0.113	0.030	-0.020
Greek	0.682	0.679	0.550	0.200
Italian	0.272	0.353	0.140	-0.020
Jewish	0.703	0.756	0.250	0.080
Northern European	0.082	0.118	0.110	0.060
Polish	0.651	0.570	0.600	0.090
Portuguese	0.785	0.726	0.610	0.170
South E. Asian	0.754	0.715	0.540	0.090
Southern European	0.462	0.498	0.340	0.070
Western European	-0.446	-0.357	-0.160	-0.040

Like the dissimilarity index, the relative concentration index moves closer to zero as the number of spatial units approaches one. This index ranges from -1 to +1, with zero suggesting an equal concentration of group x and y; +1 suggesting that x's concentration exceeds y's to the maximum extent possible; and -1 suggesting the reverse. In Table 3.31, the concentration of the East Indian group at the

EA level exceeded the concentration of the English group by nearly 72%. According to Denton (1994), a value of 0.7 is considered to be segregated under stringent criteria and 0.6 is considered high under moderate criteria. The Black, Chinese, Portuguese, East Indian and Jewish groups all experienced high values at the EA level. The Western European group was the only group (at the EA level) that showed a negative value, which meant that the Western European group was the only group that was less concentrated than the English group y. In other words, the Western European group was the only group that had a higher share of urban space than the English.

3.1.3.2. Significance Test

The Friedman test was again used to test whether the values of the RCO index were significantly different.

Table 3.32: Friedman Two-Way Anova Test comparing RCO values.

Number of cases read: 20 Number of cases listed: 20

- - - - - Friedman Two-Way Anova

Mean Rank	Variable
3.28	EA
2.28	FED
1.38	CSD
3.08	CT

Cases	Chi-Square	D.F.	Significance
20	26.9700	3	.0000

Because the Friedman test (Table 3.32) returned a low significance value, the null hypothesis (H_0) was rejected. This meant that the values of the RCO index changed significantly when they were calculated at four levels of aggregation. However, although the Friedman test returned a low significance value, some of the mean rank values looked very similar. To explore this further, a correlation test was performed. As revealed in Table 3.33, the CT and EA samples were highly correlated (.99). The FED was also highly correlated with the CT and EA levels of aggregation; however, the CSD level was not as correlated with the other levels (<.6). This meant that the CSD boundaries did not correspond with the underlying data. What was interesting was that there was not much of a difference between the number of spatial units between the FED (37) and CSD (27) level of aggregation. However, according to these results, much of the variance was lost when using CSD boundaries, while the FED boundaries kept much of the variance. Therefore, ethnic enclaves in Toronto followed FED boundaries more than CSD boundaries. Thus, CSD's were not meaningful units in terms of ethnic origin data.

Table 3.33: Correlation test of RCO index values at standard levels of aggregation.

	EA	FED	CSD	CT
EA	1.0000 P= .	.9285 P= .000	.5102 P= .022	.9898 P= .000
FED	.9285 P= .000	1.0000 P= .	.5514 P= .012	.9020 P= .000
CSD	.5102 P= .022	.5514 P= .012	1.0000 P= .	.5620 P= .010
CT	.9898 P= .000	.9020 P= .000	.5620 P= .010	1.0000 P= .

3.1.3.3. Segregation Ranks

The RCO ranks were also compared individually. According to Table 3.34, the RCO rank for most of the groups stayed relatively the same at different levels of aggregation. This was expected because of the high correlation between the different levels of aggregation. Also, because the EA and CT levels were highly correlated (.9898), one would assume that the EA level would not tell us more than the CT level. However, by using both the EA and CT levels of aggregation, different spatial patterns were revealed. For example, by using EA level data, the Black group was shown to live primarily in small spatial units that were scattered throughout the urban areas of the study. However, by using CT level data, this distribution pattern was partially lost.

Table 3.34: RCO Segregation Ranking.

Rank	Ethnic Group (EA)	Ethnic Group (CT)	Ethnic Group (FED)	Ethnic Group (CSD)
20	Western European	Western European	Western European	East Indian
19	German	German	Canadian	Western European
18	English	English	English	Canadian
17	Canadian	Canadian	German	German
16	Northern European	Northern European	Northern European	Italian
15	Italian	Italian	Italian	Arab (Lebanese)
14	Southern European	French	Jewish	English
13	Aboriginal	Eastern European	French	Filipino
12	French	Southern European	Southern European	French
11	Eastern European	Aboriginal	Aboriginal	Black
10	Polish	Polish	East Indian	Northern European
9	Greek	East Indian	Filipino	Southern European
8	Filipino	Filipino	Chinese	Jewish
7	Jewish	Greek	Eastern European	South E. Asian
6	East Indian	Arab (Lebanese)	South E. Asian	Polish
5	South E. Asian	South E. Asian	Greek	Eastern European
4	Arab (Lebanese)	Black	Polish	Chinese
3	Chinese	Portuguese	Arab (Lebanese)	Aboriginal
2	Portuguese	Chinese	Black	Portuguese
1	Black	Jewish	Portuguese	Greek

According to Table 3.34, the most noticeable changes in rank occurred with the Jewish, Blacks, and East Indians: the ranking of the Jewish group varied substantially, from 1st (CT) to 14th (FED); the ranking of the Black group went from 1st (EA) to 11th (CSD); and the ranking of the East Indian group went from 6th (EA) to 20th (CSD). Notaby, the CSD level revealed a different pattern than the other levels of aggregation. For example, at the EA, CT and FED levels of aggregation, the Western Europeans and Canadians were the only groups that were less concentrated than the English; at the CSD level, the East Indians, Western Europeans, Canadians, Italians, and Arabs were less concentrated than the English. The reason that many groups were less concentrated than the English at the CSD level was because these groups had a high proportion of their population in large peripheral areas and in other large CSDs. Although the English group was strongly represented in peripheral areas, they had a stronger representation in Metropolitan Toronto. Therefore, their population within the city was given greater weight.

3.1.3.4. Residential Calculations

Table 3.35: RCO index values calculated with residential area only.

GROUP	EA	CT	FED	CSD
Aboriginal	0.357	0.280	0.250	-0.065
Arab	0.393	0.279	0.211	-0.373
Black	0.429	0.317	0.269	-0.260
Canadian	0.000	-0.008	-0.020	-0.007
Chinese	0.250	0.262	0.178	-0.253
Eastern European	0.286	0.244	0.253	-0.150
East Indian	0.250	0.185	0.072	-0.268
English	0.000	0.000	0.000	0.000
Filipino	0.250	0.190	0.105	-0.256
French	0.214	0.168	0.112	-0.072
German	-0.071	-0.022	0.007	-0.039
Greek	0.179	0.241	0.254	-0.096
Italian	-0.215	-0.015	-0.230	-0.112
Jewish	0.107	0.155	-0.200	-0.179
Northern European	0.000	0.022	0.012	-0.026
Polish	0.393	0.300	0.280	-0.190
Portuguese	0.393	0.398	0.341	-0.058
South East Asian	0.286	0.267	0.190	-0.257
Southern European	0.036	0.149	0.004	-0.113
Western European	-0.107	-0.088	-0.071	0.017

The RCO index was also calculated by using residential area only. The recalculated RCO index values for the four levels of aggregation (Table 3.35) were all less than the original values (Table 3.31). For example, the highest value using residential area was 0.429 (Black - EA), compared to 0.810 (Black - EA) from the original calculations. Also, there were more negative values in the new calculations. The main reason for the difference in values was the reduction in area in the rural units. As mentioned previously, the RCO index was calculated by ordering the spatial units by area and then comparing the minimum with the maximum amount of area a group could occupy. Although the cumulative proportion of the area was used, the absolute area itself was never used in the calculations. What made a difference was the order and population of the spatial units. When some of the larger spatial units were drastically reduced, the order of the spatial units changed dramatically. Therefore, the large spatial units were given a different weight than in the original calculations. Also, since the RCO index was relative, the value of the index reflected how concentrated a group was in comparison to the core group (group y). In addition, when large rural spatial units were drastically reduced, it affected groups with higher populations in these units, e.g. the English and Western Europeans. Consequently, all the groups had lower values when only residential area was used because the relative difference between the English and the other groups was not as extreme as in the original calculations. At the CSD level, the new residential values were dramatically different; all the groups had negative values except for the Western Europeans. To explore this further, the original ranks were compared with the new residential ranks.

Table 3.36: Comparing Original RCO Ranks with Residential RCO Ranks

RANK	Residential Rank EA	Original Rank EA	Residential Rank CT	Original Rank CT	Residential Rank FED	Original Rank FED	Residential Rank CSD	Original Rank CSD
1	Black	Black	Portuguese	Jewish	Portuguese	Arab /Lebanese	Western European	Aboriginal
2	Portuguese	Portuguese	Black	Chinese	Polish	Black	English	Arab /Lebanese
3	Polish	Chinese	Polish	Portuguese	Black	Portuguese	Canadian	Black
4	Arab /Lebanese	Arab /Lebanese	Arab /Lebanese	Black	Eastern European	Polish	Northern European	Canadian
5	Aboriginal	South E. Asian	Aboriginal	South E. Asian	Aboriginal	Greek	German	Chinese
6	South E. Asian	East Indian	South E. Asian	Arab /Lebanese	Greek	South E. Asian	Portugu-ese	East Indian
7	Chinese	Jewish	Eastern European	Greek	Chinese	Eastern European	French	Eastern European
8	East Indian	Filipino	Chinese	Filipino	Arab /Lebanese	Chinese	Aboriginal	English
9	Eastern European	Greek	Greek	East Indian	South E. Asian	Filipino	Greek	Filipino
10	Filipino	Polish	Filipino	Polish	French	East Indian	Italian	French
11	French	Eastern European	East Indian	Aboriginal	Filipino	Aboriginal	Southern European	German
12	Greek	French	French	Southern European	East Indian	Southern European	Eastern European	Greek
13	Jewish	Aboriginal	Jewish	Eastern European	Northern European	French	Jewish	Italian
14	Southern European	Southern European	Southern European	French	German	Jewish	Polish	Jewish
15	Canadian	Italian	Northern European	Italian	Southern European	Italian	Filipino	Northern European
16	English	Northern European	English	Northern European	English	Northern European	Black	Polish
17	Northern European	Canadian	Canadian	Canadian	Canadian	German	South E. Asian	Portuguese
18	German	English	Italian	English	Western European	English	Chinese	South E. Asian
19	Western European	German	German	German	Jewish	Canadian	East Indian	Southern European
20	Italian	Western European	Western European	Western European	Italian	Western European	Arab /Lebanese	Western European

Table 3.36 compared the original RCO ranks with the residential RCO ranks. The most noticeable change was in the residential RCO ranks at the CSD level, where all the ranks were the reverse of all the other aggregation levels. For example, the Western European group moved from being the least concentrated group at the EA, CT and FED levels, to being the most concentrated group at the

CSD level. The Western European group was the most concentrated at the CSD level because the CSD boundaries of the rural spatial units had the least amount of area and the highest proportion of Western Europeans. At the EA, CT, and FED levels of aggregation, most of the population of most of the groups was in small spatial units within Metropolitan Toronto. Also, at these levels, most of the urban spatial units were smaller than the rural spatial units. Therefore, the difference at the CSD level had to do with changes in the spatial unit boundaries.

Although most of the residential ranks stayed the same at the EA, CT and FED levels of aggregation, some of the ranks changed slightly. For example, the relative position of the Polish group became more concentrated with the residential calculations. This higher concentration with residential calculations indicated that the spatial units with a high proportion of Polish ethnics had their area greatly reduced when non-residential area was removed. Notably, most of the population of the Polish group was within Southern Etobicoke.

3.1.4. Centralization

"Centralization refers to nearness to the Central Business District [CBD] or downtown area" (Denton, 1994, p. 54). The centralization dimension rests upon the premise that "in most larger urban areas, much of the area surrounding the downtown area is old, rundown, and creating what Park and Burgess (1925) classed the zone of transition" (Denton, 1994, p. 54). However, this premise is not fully supported in Toronto. Unlike most American cities, many minorities in Toronto's CMA live outside the city centre. For example, in Toronto there are twice as many people of single Black origin that live in North York and Scarborough than in the city of Toronto. Part of the reason for this is that cheaper rental accommodation is often found in the older suburbs. Also, there are forces such as gentrification which counter this idea. Despite conceptual problems, however, the centralization dimension is still considered relevant.

3.1.4.1. Index Values

The centralization dimension has been measured by using the absolute centralization index (ACE index). Table 3.41 shows the ACE index values for various groups within Toronto's CMA.

Table 3.41: **ACE** index for Toronto's CMA (1991) at four standard levels of aggregation

Ethnic Group	ea (5370	ct (812)	fed (37)	csd (27)
Aboriginal	0.00	0.00	0.00	0.00
Arab (Lebanese)	0.78	0.78	0.77	0.77
Black	0.78	0.78	0.77	0.77
British	0.61	0.61	0.59	0.59
Canadian	0.57	0.57	0.55	0.55
Chinese	0.82	0.82	0.80	0.80
Eastern European	0.78	0.78	0.75	0.75
East Indian	0.74	0.74	0.72	0.72
English	0.59	0.60	0.57	0.57
Filipino	0.78	0.78	0.75	0.75
French	0.68	0.68	0.66	0.66
German	0.62	0.62	0.60	0.60
Greek	0.83	0.83	0.81	0.81
Italian	0.76	0.76	0.74	0.74
Jewish	0.87	0.86	0.84	0.84
Northern	0.66	0.66	0.63	0.63
Polish	0.80	0.80	0.77	0.77
Portuguese	0.80	0.80	0.78	0.78
South E. Asian	0.82	0.82	0.79	0.79
Southern	0.78	0.78	0.76	0.76
Western European	0.56	0.56	0.54	0.54

The ACE index ranges from +1 to -1, with 0 representing uniform distribution across the city, +1 representing a tendency for group x to live close to the Central Business District (CBD), and -1 representing a tendency for group x to live in outlying areas. For example, the ACE index value of 0.8 (EA) for the Portuguese group meant that 80% of the Portuguese group had to move further from the CBD in order to achieve uniform distribution around the CBD. Also, the ACE index value of 0.6 for the German group suggested that 60% of the German group had to move further from the CBD in order to achieve uniform distribution around the CBD. Notably, the ACE index gives only a relative description: closer to or further from. The ACE index does not tell how close a group is to the CBD. In addition, this index does not tell us anything about where a group is coming from or going to in order to achieve uniform distribution. Rather it is a measure to be used for comparison with other groups: group A has a higher ACE index value than group B; therefore, group A is more centrally located than group B. Accordingly, Table 3.41 revealed that all of the groups had positive values using the ACE index. This made sense since the population of most groups was within Metropolitan Toronto. The only exception was the aboriginal group with a value of zero. This was in spite of the fact that an Indian reserve located on Georgina island was not included in the calculations. If it was, the Aboriginal ACE index may have had a negative value. Although the rest of the groups had high positive values, the higher values of some of the groups suggested that these groups were more centrally located than others. The highest values came from the Jewish, Greeks, Chinese, South East Asians, Polish and Portuguese. In contrast, the French, Northern Europeans, Germans, British, English, Canadians, Western Europeans and Aboriginals had the lowest values. Generally, the white ethnic groups had higher centralization values than the visible minorities. This contrasts dramatically with the pattern that exists in most American cities.

In the United States there is usually a direct relationship between closeness to the CBD and economic standing: economic standing increases the further you move from the CBD. In many third-world cities this relationship is opposite: economic standing increases the closer you move to the CBD. However, Canadian cities, in general, and Toronto, in particular, do not have the same direct relationship between closeness to the CBD and economic standing. Despite this, the measurement of centralization is useful in Toronto as a suburbanization index. In Toronto, for example, the core ethnic groups had the lowest ACE values. This makes sense since living outside Metropolitan Toronto in the newer suburbs often requires greater wealth and mobility. Significantly, the intermediate groups were the groups with the lowest economic standing. Economic standing seemed to be more related to access to public transportation than to closeness to the CBD. In Toronto, public housing has been purposely decentralized and scattered to try to counter some of the negative effects of segregation that have been found in many American cities. Importantly, there is very little ethnic representation outside Metropolitan Toronto, Brampton, and Mississauga. Perhaps the centralization dimension would have been more useful if it measured distance to public transportation than closeness to the CBD.

Another problem with the centralization dimension is the assumption that all people within an ethnic group have the same economic standing. In Toronto, for example, the Chinese group exhibits a polar economic distribution with the wealthy Chinese ethnics living mainly in the newer suburbs and the poor Chinese ethnics living mainly in Chinatown (centred at Dundas and Spadina). Chinatown is still the main reception area for poor Chinese immigrants. Although the Chinese example seems to suggest that the immigrants downtown are poorer than their suburban counterparts, this is not necessarily true. In fact, "the majority of these immigrants (excluding refugees) are not low-income, even though their incomes are below the CMA average (de Silva, 1992). Moreover, most of the immigrants with higher incomes now also go directly to the new suburbs rather than the traditional immigrant reception areas in the central city, while lower income immigrants are increasingly drawn to the mature suburbs where low rent housing (both public and private) is more widely available. This divergence of immigrant flows, in turn, has contributed to the suburbanization of wealth and growing disparity between the inner suburban municipalities and the newer suburbs" (Bourne, 1996, p. 12).

3.1.4.2. Significance Test

According to Table 3.41, there was virtually no difference between the various levels of aggregation when using the ACE index. The CSD level (27 spatial units) achieved nearly the same values as the EA level (5,370 spatial units). Therefore, there was no advantage to using EA level data to calculate this index. The Friedman test was not used here because the four sets of data were essentially the same. Instead, a correlation test was calculated, which showed that the lowest correlation between aggregation levels was 0.9993 (P=0.000). Notably, the EA and CT levels and the FED and CSD levels were nearly identical. One reason that there was little or no change between aggregation levels was that the distance remained constant. Even when the data was aggregated, the order still remained the same. Because the distance did not change between different levels of aggregation, the cumulative proportions cancelled out the increased detail of the smaller spatial units. Thus aggregation had little effect upon the index because cumulative proportions were used instead of absolute values. Although the correlation between the aggregation levels was not perfect, there was some reduction in the ACE index when the number of spatial units was reduced.

3.1.4.3. Segregation Ranks

To further understand the ACE index and how it behaves at different levels of aggregation, it was felt worthwhile to look at the minimum number of spatial units. The minimum number of spatial units of the ACE index is two because there is a lag in calculation (x-1). This contrasts with the other indices where the minimum number is one spatial unit. If we take the example of two spatial units, one being Metropolitan Toronto and the other the 905 belt[iii], the equation would simplify into the following: ACE = (the proportion of group x in Metropolitan Toronto) - (the proportion of area outside Metropolitan Toronto). The Jewish group, for example, would have an index value of 0.585, which is much lower than the values of the Jewish group in Table 3.41. Although the values go down when the number of spatial units is reduced, this is not significant until the number of spatial units becomes very small. Breaking the index down into two spatial units reveals that the choice of aggregation boundary is more important when there are fewer spatial units. The index value would be different if the spatial units were broken down into the city of Toronto and the rest of the study area. Why then was there virtually no change between the four levels of aggregation? One reason for this is the way that Census Canada aggregates its data. Although attempts are not always perfect, Census Canada tries to make the spatial units as homogeneous as possible at each level of aggregation. With the ACE index, the choice of aggregation boundary becomes less important when the number of spatial units increases.

[iii] The 905 belt refers to places outside metropolitan Toronto and have a 905 area code.

3.1.4.4. Residential Calculations

The ACE index was also calculated using residential land only. Only the EA level was calculated because the original calculations showed that there was virtually no difference between aggregation levels for the ACE index. Table 3.42 showed the new ACE index values using residential land only. The values in Table 3.42, however, were considerably lower than the original calculations in Table 3.41. For example, the highest value in Table 3.42, 0.534 (Jewish [EA]), can be compared to the highest value in Table 3.41, 0.87 (Jewish [EA]). Because the ACE index was calculated by ordering the spatial units by distance, changing the area of the spatial units did not change the distance. What did change was the amount of area in the study, especially in rural regions, which affected the cumulative proportion of the area. As the area of the largest units was greatly reduced, this reduced the variance in area between the spatial units. Since a cumulative proportion was used, the reduction in variance increased the proportion of total area in the centre of the city and thus gave central units more weight than in the original calculations. Therefore, the reduction in area, especially in peripheral areas, gave the illusion of a more centralized distribution around the CBD, even though the distance had not changed.

Table 3.42: ACE index values calculated with residential area only

Ethnic Group	ACE (EA) Residential
Aboriginal	0.000
Arab	0.332
Black	0.332
British	0.159
Canadian	0.100
Chinese	0.433
Eastern European	0.414
East Indian	0.222
English	0.143
Filipino	0.340
French	0.249
German	0.169
Greek	0.505
Italian	0.306
Jewish	0.534
Northern European	0.239
Polish	0.428
Portuguese	0.485
South East Asian	0.425
Southern European	0.375
Western European	0.088

Table 3.43: Comparing original ACE ranks with residential ACE ranks

Rank	Original Area	Residential Area
1	Jewish	Jewish
2	Greek	Greek
3	Chinese	Portuguese
4	South East Asian	Chinese
5	Polish	Polish
6	Portuguese	South East Asian
7	Arab	Eastern European
8	Black	Southern European
9	Eastern European	Filipino
10	Filipino	Arab
11	Southern European	Black
12	Italian	Italian
13	East Indian	French
14	French	Northern European
15	Northern European	East Indian
16	German	German
17	British	British
18	English	English
19	Canadian	Canadian
20	Western European	Western European
21	Aboriginal	Aboriginal

The question remained as to whether the ranking of the ethnic groups would stay the same when the index was recalculated using residential area only. Table 3.43 compared the original ranks with the new residential ranks of the ACE index. Accordingly, the ranks of most of the groups stayed either the same or moved only slightly, e.g. moved one to four ranks. Thus, when the ACE index was recalculated using residential area only, the absolute values of all the indices were reduced, however, the relative standing of most of the groups stayed basically the same.

3.1.5. Clustering

The clustering dimension addresses the "checkerboard problem." "Segregation is worse if all the red squares are on one side and all black ones on the other, than if the individually segregated neighbourhoods are mixed up like the squares on a checkerboard" (Denton, 1994, p. 54). It is generally accepted that fewer, larger clusters are worse than smaller clusters that are scattered throughout the city. Clustering is different from concentration in that concentration measures the density of each spatial unit in isolation. Clustering measures a proportion of the population of a spatial unit compared to a proportion of the population of neighbouring spatial units. The clustering dimension is the only dimension that takes into account neighbouring spatial units. Autocorrelation is similar to clustering, which is measured by the Moran's I and Geary's C statistic.

In order to measure clustering or autocorrelation, a contiguity matrix was needed. A contiguity matrix is a matrix identifying whether one spatial unit borders upon another by using an element known as c_{ij}: c_{ij} equals one if units i and j are contiguous or share a border, and zero for all other cases. To measure clustering, the spatial proximity index was used. However, the spatial proximity index used a distance delay function which approximated a contiguity matrix. Element c_{ij} was estimated by an inverse distance function (see Chapter 2), where the further spatial unit i was from spatial unit j, c_{ij} approached zero; conversely, the closer spatial unit i was to spatial unit j, the closer c_{ij} moved towards one. A value of one would occur if spatial units i and j were the same, because in this case, there would be no distance between the two spatial units.

3.1.5.1. Index Values

Table 3.51: Spatial Proximity index for Toronto's CMA (1991) at EA and CT levels of aggregation

Ethnic Group	ea (5370) English(y)	ea (5370) British(y)	ct (812) English(y)	ct (812) British(y)
Aboriginal	0.123	0.179	0.144	0.206
Arab (Lebanese)	0.126	0181	0.147	0.209
Black	0.167	0.223	0 .202	0.264
Canadian	0.179	0.235	0.213	0.276
Chinese	0.235	0.290	0.279	0.342
Eastern European	0.180	0.236	0.208	0.271
East Indian	0.171	0.226	0.207	0.270
English	0.239	0.294	0.279	0 .341
Filipino	0.146	0.202	0.175	0.238
French	0.135	0.191	0.155	0.218
German	0.137	0.193	0.158	0.221
Greek	0.149	0.204	0.177	0.239
Italian	0.236	0.292	0.327	0.390
Jewish	0.208	0.263	0.389	0.452
Northern European	0.124	0.180	0.145	0.207
Polish	0.158	0.213	0.187	0.250
Portuguese	0.227	0.283	0.275	0.337
South East Asian	0.267	0.322	0.303	0.366
Southern European	0.318	0.374	0.389	0.452
Western European	0.144	0.200	0.168	0.231

Table 3.51 illustrates the spatial proximity index values for the various ethnic groups (with the English and British as group y) at the enumeration and census tract levels of aggregation. Although all of the groups had low index scores, some of the groups had higher scores than others. The highest values came from the Jewish, Southern Europeans, Portuguese and Chinese; these relative findings are

consistent with other findings on these groups. Only the EA and CT aggregation levels were calculated with a 1 km buffer. It was felt that the FED and CSD units were too large in area and too few in number to be appropriate for this index.

The spatial proximity index, however, had several problems:

i. heavy computational requirements;

ii. ambiguity of distance measurement and choice of measurement used;

iii. confusing relative relationship;

iv. sensitivity of index to group size.

(i) The computational requirements of the spatial proximity index were heavy because of the large amount of data that was produced in making a contiguity or distance matrix. For example, in Toronto's CMA, there were 5,370 enumeration areas, which created a distance matrix table of approximately 18,836,900. A reduced 1 km calculation still resulted in a table of approximately 152,000. Notably, with continued advancements in computer technology, large files such as this will become less of a problem.

(ii) Another problem was the ambiguity of the distance measurement. In this study, a 1km buffer was used to reduce the size of the file into a more manageable form. This was not considered to be much of a problem because the inverse distance value drastically reduced after the distance measurement went above one. More problematic, however, was the choice of measurement. Because the spatial proximity index was dependent upon the value of the distance measurement, the choice of measurement was very important. The use of kilometers or miles produced a dramatic difference. Table 3.52 illustrates the different c_{ij} values at different distances and different units of measure.

Table 3.52: Example C_{ij} values for various distances and units of measure.

Metres	C_{ij} (m)	Kilometres	C_{ij} (km)	Miles	C_{ij} (mi)
1	.368	.001	.999	.000625	.999
1000	0	1	.368	.625	.535
2000	0	2	.135	1.25	.287
5000	0	5	.0067	3.125	.044
10000	0	10	.000	6.25	.002
1600	0	1.6	.202	1	.368
3200	0	3.2	.041	2	.135
8000	0	8	.000	5	.0067
16000	0	16	.000	10	.000

For example, a distance of one mile (1.6 km) would translate into a c_{ij} value of 0.368 if miles were used, compared to 0.202 if kilometres were used. Notably, in many studies, the unit of measure is often not indicated. For true comparison, however, the choice of measurement should be indicated and/or a standard measurement used.

(iii) The spatial proximity index is a relative index because it uses group y in its calculations. This relative relationship, however, is not made clear to the reader. According to Table 3.51, the English group (group y) had higher spatial proximity values than most of the other groups. From this, one could conclude that the English was more clustered than the other groups. However, this reflects their population size, rather than their clustering. The relative relationship is usually clearer when the value of group y represents zero.

(iv) Another problem that occurred was the sensitivity of the spatial proximity index to group size. Although group size also affected the exposure dimension, it was not considered as problematic because exposure (or isolation) is directly related to the probability of one group interacting with another. The clustering dimension, on the other hand, is more interested in proximity than in group or cluster size. When changing from EA to CT levels of aggregation, all the SP index values went up (see Table 3.51). This was the reverse of all other indices. The higher values at the CT level could be attributed to the dependence of the SP index upon population size. In addition, the values of the SP index went up when group y was changed from English to British, because the British, group y, had a larger population than the English, group y. Also, comparing the Jewish and Southern European groups illustrated how dependent the SP index was upon group size. Although the Jewish and Southern European groups had the same index values, the reasons were very different. The Jewish had a high index value because they were highly clustered, but had a smaller proportion of the total population. In contrast, the Southern Europeans were not as clustered, but had a larger proportion of the total population. Thus, the values of the SP index reflected clustering and group size. This became more apparent in Table 3.53, where several indices were calculated and compared.

To examine the spatial proximity index in more detail, the values of the index were compared with the original P_{xx} values. The P_{xx} value was the clustering component of the spatial proximity index. As indicated in Chapter 2, the P_{xx} value was the average proximity between x members. This was calculated by multiplying a proportion of group x within spatial units i and j with an inverse distance function ($c_{ij} = \exp(-d_{ij})$). Thus, the further spatial units i and j moved from each other, the smaller c_{ij} became, while the closer they moved together, c_{ij} approached 1. By using a proportion of group x rather than the total population, the problem of group size was diminished. Notably, P_{xx} values do not reflect group size. Therefore, if group size is reflected in the index, it is because of other components. Three other indices using P_{xx} as a clustering component were also calculated and compared to the spatial proximity index. Although these indices also used P_{xx} values, they left a different impression. Table

3.53 shows four indices that use P_{xx} as their main clustering component: the spatial proximity index, the RCL index, and two hybrids of the RCL index (RCL_new and RC).

Table 3.53: Clustering Indices for Toronto's CMA (1991), by Census Tract

Ethnic Group	**RCL**	**RCL_new**	**RC**	**SP English(y)**	**P_{xx}**
Aboriginal	1.709	-0.969	0.631	0.144	0.007
Arab (Lebanese)	2.168	-0.949	0.684	0.147	0.008
Black	0.847	-0.554	0.459	0 .202	0.005
Canadian	0.033	-0.471	0.032	0.213	0.003
Chinese	1.249	0.004	0.555	0.279	0.006
Eastern European	0.889	-0.507	0.471	0.208	0.005
East Indian	0.784	-0.515	0.439	0.207	0.004
English	0	0	0.000	0.279	0.002
Filipino	1.002	-0.745	0.500	0.175	0.005
French	0.169	-0.885	0.145	0.155	0.003
German	0.055	-0.863	0.053	0.158	0.003
Greek	1.214	-0.732	0.548	0.177	0.005
Italian	1.247	0.347	0.555	0.327	0.006
Jewish	7.117	0.792	0.877	0.389	0.02
Northern European	0.544	-0.962	0.352	0.145	0.004
Polish	1.465	-0.657	0.594	0.187	0.006
Portuguese	3.059	-0.031	0.754	0.275	0.01
South East Asian	0.718	0.172	0.418	0.303	0.004
Southern European	0.658	0.789	0.397	0.389	0.004
Western European	0.087	-0.794	0.080	0.168	0.003

Spatial Proximity Index:

Although the spatial proximity index should ideally reflect P_{xx} values, it often reflects group size. As mentioned previously, although the Jewish and Southern European groups had the same SP values, their P_{xx} values were very different, e.g. Jewish (0.020) and Southern Europeans (0.004). The Jewish group had P_{xx} values that were five times higher than the Southern European group; however, their similar SP values suggested that they were equally clustered. The main reason that their SP values were the same was because of group size. Another example that illustrated the problem of group size was the English group. This group had the lowest P_{xx} value (.002); however, it also had a reasonably high SP value (their SP value equalled that of the Chinese). Thus, the values of the spatial proximity index reflected population size and spatial clustering.

RCL index:

The RCL index was measured by $RCL = (P_{XX}/P_{YY}) - 1$, where P_{XX} was the average proximity between group x members, and P_{YY} was the average proximity between group y members. Therefore, the calculation of P_{YY} values involved substituting group y values for group x values. As indicated in Table 3.53, the RCL value of the English group was zero because it divided itself, minus one. The RCL index value was 7.117 for the Jewish and 0.087 for the Southern Europeans. Because these values reflected P_{XX} values, they were more reflective of the clustering dimension. The main argument against the RCL index was that it had no theoretical maximum. As discussed previously, this lack of standardization is problematic with regard to comparative studies. The two remaining indices, the RCL_new and the RC, tried to accommodate this problem.

RCL_new index:

The RCL_new index is a hybrid of the RCL and spatial proximity indices. It was calculated by using the following equation: RCL_new = ((x * pxx) / (y_eng * pyy_eng)) - 1, where x is the population of group x for the entire area, and y_eng is the population of the English, group y. The relative relationship is clearer with the RCL_new index because, like the RCL index, the value of the English group was 0. Therefore, if a group was equally as clustered as the English group, it would also have a value of 0. However, group size was still a factor with the RCL_new index. Thus the Jewish and Southern European groups still had the highest values.

RC index:

The relative clustering index (RC index) is a mutated version of the RCL index. The RC index was calculated by using the following formula: $RC = 1 - (P_{YY}/P_{XX})$. Like the RCL index, the RC index did not have the problem of group size that was experienced by the SP and RCL_new indices. The RC index measured how much more group x was clustered than group y (the core group). For example, the value of 0.9 for the Jewish group suggested that this group was 10 times more clustered than the English group: 0.5 would mean that group x was 2 times more clustered than group y. A standard value of 0.7 would mean that a group was segregated if it was approximately 3 times more clustered than group y. Overall, the RC index had several advantages. First, the index was easy to understand. Second, the index had a theoretical maximum of 1 (unlike the RCL index). Although there was no theoretical minimum (maximum negative value), this was not felt to be as important because most groups were considered to be more clustered than the core group. If a group was less clustered, it would only be slightly less clustered. However, if several groups were significantly less clustered, the core group should be re-evaluated. With Toronto, the English group had the lowest value so this was not an issue. Group y could also be assigned by choosing the group with the lowest P_{XX} value. This would result in no negative index values.

Although the spatial proximity index had many advantages, it also had disadvantages. For example, the spatial proximity index was affected by group size and lacked a clear relative relationship. An alternative to the spatial proximity index was the RCL index. The RCL index was simpler and portrayed a more meaningful relationship relative to the core group. However, the RCL index had no theoretical maximum, which made it harder to read and compare. To try to overcome the problems incurred by these indices, two alternate indices were recommended, the RCL_new and the RC indices. The RCL_new index is a hybrid of the spatial proximity index and the RCL index. The RCL_new index helped to make the relative nature of the index easier to read and understand; however, the problem of group size still existed. The RC index is a hybrid of the RCL index alone. The RCL index had the advantage of a theoretical maximum, an easily understood relative relationship, and an easily interpreted index value.

3.1.5.2. Significance Test

The Wilcoxon Matched-Pairs Signed-Ranks Test was used instead of the Friedman test because only two aggregation levels were used. The Wilcoxon test was read in the same way as the Friedman test.

Table 3.54: Wilcoxon Matched-Pairs Signed-Ranks Test comparing Spatial Proximity index values

```
   CT
with EA

  Mean Rank        Cases

   10.50  20  - Ranks (EA LT CT)
     .00   0  + Ranks (EA GT CT)
           0    Ties (EA EQ CT)
          --
          20    Total

   Z =  -3.924        2-Tailed P =  .0001
```

Since the Wilcoxon test returned a low significance value (.0001), the null hypothesis was rejected. The values of the SP index at the EA and CT levels of aggregation were therefore significantly different.

3.1.5.3. Segregation Ranks

The SP index ranks were also compared individually (see Table 3.55). Although the absolute values of the SP index changed significantly, the ranks either did not change or changed only slightly. The Jewish group showed the greatest change, from a rank of 7 at the EA level, to a rank of 1 (tied) at the CT level. Similar to the changes in absolute SP values, the change in rank can be attributed to the P_{xx} values since the population of a group remained constant at different levels of aggregation. Therefore, the P_{xx} values for the Jewish group increased more dramatically than the other groups when moving from EA to CT levels of aggregation.

Table 3.55: Spatial Proximity Index Ranks at EA and CT Levels of Aggregation

Rank	EA	CT
1	Southern European	[redacted]
2	South East Asian	[redacted]
3	English	Italian
4	Italian	[redacted]
5	Chinese	[redacted]
6	Portuguese	Chinese
7	Jewish	Portuguese
8	Eastern European	Canadian
9	Canadian	Eastern European
10	East Indian	East Indian
11	Black	Black
12	Polish	Polish
13	Greek	Greek
14	Filipino	Filipino
15	Western European	Western European
16	German	German
17	French	French
18	Arab (Lebanese)	Arab (Lebanese)
19	Northern European	Northern European
20	Aboriginal	Aboriginal

3.2. Segregation Dimensions Viewed Together

The five dimensions of segregation were also looked at in combination. The dimensions were viewed (a) in terms of hypersegregation, and (b) in terms of overall ranking.

3.2.1. Hypersegregation

Hypersegregation has become a popular term in segregation studies since the 1980s. "Following the 1980 U.S. census, the term hypersegregation became part of American public parlance" (Denton, 1994, p. 49). Testing for hypersegregation in Toronto provides a broader and multilayered approach which is useful for comparative studies. Also, it is useful because hypersegregation has not been tested before in Toronto. The term hypersegregation refers to high index values on at least four of the five dimensions of segregation. Two sets of criteria that are used to test for hypersegregation are stringent and less stringent.

Table 3.61: Stringent and less stringent criteria for hypersegregation, by dimension.

Dimension	Stringent Criteria	Less Stringent Criteria
Evenness	>0.6	>0.6
Isolation	>0.7	>0.6
Concentration	>0.7	>0.6
Centralization	>0.8	>0.6
Clustering	>0.6	>0.6

Figure 3.61 lists the ethnic groups that met the criteria in each dimension. The brackets following the ethnic groups refer to the aggregation levels where the criteria was met.

Table 3.61: Ethnic groups that met the Hypersegregation criteria (1991)

With the evenness dimension, the following groups would be labelled as segregated:

- *Jewish (EA, CT, FED)*
- *Poruguese (EA, CT)*
- *Arab (EA, CT, FED)*
- *Aborginal (EA, CT)*
- *Northern European (EA)*

With the concentration dimension, the following groups would be labeled as segregated:
>0.7
- *Black (EA, CT, FED,CSD)*
- *Jewish (EA, CT, FED)*
- *Portuguese (EA, CT, FED)*
- *Chinese (EA, CT, FED)*
- *Southeast Asian (EA, CT, FED)*

>0.6
- *Arab (Lebanese) (CT)*
- *Filipino (EA, CT)*
- *Greek (EA, CT)*

With the centralization dimension, the following groups would be labeled as segregated:
>0.8
- *Jewish (EA, CT, FED,CSD)*
- *Greek (EA, CT, FED,CSD)*
- *Chinese (EA, CT, FED,CSD)*
- *Portuguese (EA, CT, FED,CSD)*
- *Polish (EA, CT, FED,CSD)*

>0.6
- *Aboriginal (EA, CT, FED,CSD)*
- *Arab (Lebanese) (EA, CT, FED,CSD)*
- *Black (EA, CT, FED,CSD)*
- *Canadian (EA, CT, FED,CSD)*
- *Eastern European (EA, CT, FED,CSD)*
- *East Indian (EA, CT, FED,CSD)*
- *English (EA, CT, FED,CSD)*
- *Filipino (EA, CT, FED,CSD)*
- *French (EA, CT, FED,CSD)*
- *German (EA, CT, FED,CSD)*
- *Italian (EA, CT, FED,CSD)*
- *Northern European (EA, CT, FED,CSD)*
- *Southeast Asian (EA, CT, FED,CSD)*
- *Southern European (EA, CT, FED,CSD)*
- *Western European (EA, CT, FED,CSD)*

Although some of the groups met the single criteria for segregation, none of the groups met the stringent and less stringent criteria for hypersegregation. The main reason for this was that none of the groups had high scores on either the exposure or clustering dimensions. Also, the hypersegregation criteria was not appropriate for the exposure dimension because the higher the value of the interaction index, the higher the exposure to the core group (see Exposure section). The centralization dimension, in contrast, showed that although most of the groups met the less stringent criteria, only a few met the more stringent criteria. This can mainly be attributed to the scale of the study. Because the scale was Toronto's CMA, most of the population was centralized within Metropolitan Toronto and surrounding areas. If the scale of the study had been restricted to Metropolitan Toronto, the index values would have been lower and varied. Overall, the evenness and concentration dimensions had a higher range of index values.

It should be noted that most of the hypersegregation studies have been done on race within America. The smaller number of groups in a racial study creates larger group sizes. This automatically increases the values in the exposure and clustering dimensions. Therefore, the hypersegregation criteria seems to be more appropriate for the study of race than for ethnicity.

3.2.2. Overall Segregation Ranking

A look at the relative standing of the groups was thought to be useful because the clustering and exposure dimensions were artificially lowered because of the small population of most of the groups. Table 3.62 shows the average rank for each dimension across all aggregation levels and the average rank for all the dimensions combined[iv]. Overall, the most segregated group was the Jewish, followed by the Portuguese, Chinese, South East Asians, and Southern Europeans. These groups (excluding the Jewish) were the newest immigrants (post 1960). The least segregated, and therefore considered to be the most integrated, were the Northern Europeans, Western Europeans and Canadians. These groups are known as the core groups or older immigrants. Although this follows the ecological model, debate exists as to whether social economic status decreases with higher levels of integration (Massey, 1985). Also interesting and in contrast to many American cities is the fact that the Black group (also part of the newer immigrants) had an intermediate level of segregation (9th overall). The only dimension where they scored very high was the concentration (population density) dimension.

[iv] The average rank was calculated by adding the ranks of each ethnic group at each level of aggregation and then ordering them by that number.

Table 3.62: Overall segregation ranks by dimension

	All	Evenness	Isolation	Concentration	Centralization	Clustering
1	Jewish	Jewish	Jewish	Portuguese	Jewish	Southern European
2	Portuguese	Portuguese	South E. Asian	Chinese	Greek	South E. Asian
3	Chinese	Arab (Lebanese)	Italian	Black	Chinese	Italian
4	South E. Asian	Southern European	Portuguese	Greek	South E. Asian	Chinese
5	Southern European	Chinese	Chinese	South E. Asian	Polish	Portuguese
6	Arab (Lebanese)	Italian	Southern European	Arab (Lebanese)	Portuguese	Jewish
7	Greek	Polish	Black	Jewish	Arab (Lebanese)	Arab (Lebanese)
8	Italian	Aboriginal	East Indian	Polish	Black	East Indian
9	Black	Greek	Arab (Lebanese)	Eastern European	Eastern European	Canadian
10	Polish	Eastern European	Filipino	Aboriginal	Filipino	Eastern European
11	Eastern European	Black	Greek	Filipino	Southern European	Black
12	East Indian	Filipino	Polish	East Indian	Italian	Polish
13	Filipino	South E. Asian	Eastern European	Southern European	East Indian	Greek
14	Aboriginal	East Indian	Aboriginal	French	French	Filipino
15	French	Northern European	French	Northern European	Northern European	Western European
16	Northern European	Western European	Northern European	Italian	German	German
17	Canadian	Canadian	German	Canadian	Canadian	French
18	German	French	Canadian	German	Western European	Northern European
19	Western European	German	Western European	Western European	Aboriginal	Aboriginal

3.3. Summary

3.3.1. Advantages to Measuring Indices at Different Levels of Aggregation

This study found several advantages to calculating indices at different levels of aggregation:

a.. to see if the indices changed significantly at different levels of aggregation;

b. to highlight the different properties of the different indices;

c. to reveal the different spatial patterns of the different ethnic groups;

d. to illustrate the sensitivity of the indices to different levels of aggregation.

a. To see if the indices changed significantly at different levels of aggregation:

In this chapter, five indices were tested to see if they changed significantly at four standard levels of aggregation, e.g. enumeration area, census tract, federal electoral district and census subdivision. The results were that the dissimilarity, relative concentration and spatial proximity indices changed significantly. However, the xP^*y and the ACE indices did not change very much. In fact, the values of the ACE index stayed virtually the same at all levels of aggregation. It is probably more surprising that some of the indices stayed the same since information is lost and variance decreases when data is aggregated.

b. To highlight the different properties of the different indices:

Aggregating data affects each index differently. By changing the levels of aggregation, the different properties of the indices are highlighted. An index runs on a continuum between one spatial unit representing one person, to one spatial unit representing the entire study area.

With most indices, the level of segregation goes down when data is aggregated into fewer spatial units. For example, the dissimilarity index moved closer to one as the number of spatial units increased, and closer to zero as the number of spatial units decreased. Therefore, if the study area represented one spatial unit, the value of the dissimilarity index would be zero. The relative concentration index, like the dissimilarity index, moved closer to zero as the number of spatial units approached one. However, other indices behaved very differently. For example, the interaction index moved closer to the population proportion of group y (y/t) as the number of spatial units decreased. With the ACE index, the theoretical minimum is two spatial units because there is a lag in calculation (x-1). The ACE index is dependent upon the boundaries chosen when the number of spatial units is small. However, as the number of spatial units increases, the index becomes very stable. With the spatial proximity index, the index values go up as the number of spatial units decreases. Notably, the spatial proximity index was not calculated for the higher levels of aggregation (e.g. FED & CSD), because the distance measurements between the spatial units would have been too small. However, by using a hypothetical case of one spatial unit, the spatial proximity index would simplify into (X + Y)/T. An awareness of the different properties of an index is therefore useful because it makes the reader better able to discern whether index values are a product of the aggregation level or the spatial pattern that they are trying to describe.

c. To reveal the different spatial patterns of the different ethnic groups:

Calculating indices at different levels of aggregation can reveal different spatial patterns. For example, if index values reduce rapidly when the number of spatial units is reduced, the spatial distribution is more scattered. Conversely, if the index values are stable over different levels of aggregation, the spatial distribution is larger and more concentrated.

When large spatial units were used, only the ethnic groups that had large, established ethnic enclaves were identified. Furthermore, only large ethnic enclaves that conformed to the larger spatial unit boundaries were revealed. That is why some of the index values changed dramatically when FED and CSD boundaries were used, even though the difference in the number of spatial units was very little. Conversely, ethnic groups that were isolated in small spatial units were often lost when data was aggregated. By calculating indices at different levels of aggregation, the rate of change or the stability of the index values can tell us something about the spatial distribution of the different ethnic groups.

d. To illustrate the sensitivity of the indices to different levels of aggregation:

Because index values are abstract and linked to science, they are often accepted without question. However, the reader is often not aware that index values can change significantly when different spatial units are used. Therefore, demonstrating how easily index values can change, makes the reader aware that indices are open to manipulation and miscalculation.

Although the five segregation indices are widely used today, calculating indices at different levels of aggregation has not been done very much because it is too time consuming, too confusing, and the results can often be anticipated, e.g. as information is lost, variance decreases. However, despite these problems, this study has hopefully highlighted advantages to calculating indices at different levels of aggregation.

3.3.2. Advantages and Disadvantages of the Indices Themselves

Table 3.71 lists some of the advantages and disadvantages of the different dimensions and the indices used to represent them.

Table 3.71: Advantages and Disadvantages of Segregation Dimensions and Indices

Dimension	Index	Advantages	Disadvantages
Evenness	Dissimilarity Index	▸ widely used ▸ easy to calculate ▸ evenness concept easy to understand and conceptualize	▸ not aggregation consistent
Exposure	Interaction Index	▸ probability concept easy to understand and conceptualize ▸ easy to calculate	▸ index values reflect group size
	Isolation Index	▸ probability concept easy to understand and conceptualize ▸ easy to calculate	▸ index values reflect group size
Concentration	Relative Concentration Index (RCO)	▸ density measure	▸ calculation requirements heavy ▸ very time consuming ▸ not aggregation consistent
Centralization	Absolute Centralization Index (ACE)	▸ better read as a suburbanization index than as a segregation index	▸ calculation requirements heavy ▸ very time consuming ▸ questionable association of poverty with closeness to CBD
Clustering	Spatial Proximity Index	▸ conveys connectivity	▸ index values reflect group size ▸ difficult to obtain a distance matrix ▸ relative relationship not clear ▸ not aggregation consistent
	RCL	▸ conveys connectivity	▸ no theoretical maximum ▸ hard to interpret ▸ difficult to obtain a distance matrix
	RCL_new	▸ conveys connectivity	▸ difficult to obtain a distance matrix
	RC	▸ conveys connectivity	▸ difficult to obtain a distance matrix

3.3.3. Advantages to calculating the RCO and ACE indices with residential area only

In this study, the RCO and ACE indices were recalculated using residential area only. When the RCO index was recalculated using residential area only, all absolute index values were reduced. A difference was expected for the RCO index because concentration or density is directly related to area. When the ACE index [centralization] was recalculated using residential area only, all index values were also reduced. Using all of the area within the spatial unit gave the illusion of a more centralized distribution.

Was the recalculation of the ACE and RCO indices considered to be worthwhile? The new calculations were useful because they revealed that many values were artificially raised because of the large amount of unpopulated land in the rural areas of the study. However, the creation of residential spatial units was very time consuming and the relative standing for most groups stayed much the same. In addition, errors in the land use classification could be problematic. The error potential is higher in the rural areas because of the lack of detailed land use information in these areas. In addition, unless a standard land use is used, each user will have slightly different spatial units. Consequently, each user will have different index values. Thus, for comparative studies, it is preferable that all users use the same spatial units.

CHAPTER 4

CARTOGRAPHIC REPRESENTATION

4.1. Cartographic Representation

A study of ethnic residential segregation would not be complete without the use of maps. In Chapter 3, indices measured the five dimensions of segregation: evenness, exposure, concentration (density), centralization, and clustering. Maps can visually show these dimensions as well. However, maps can give more dramatic and meaningful representations by showing location/place, direction, pattern, size, shape and frequency. However, maps are also more open to manipulation and subjective interpretation.

This chapter illustrates some of the alternatives available for mapping ethnic census data. In this study, some of the issues considered when choosing maps were the visual impact of arbitrary boundaries; the suitability of different maps to population data; the suitability of different maps to different levels of scale and aggregation; the different ways of dealing with arbitrary boundaries, MAUP and areal averaging. Discussion has been broken down into two main parts. The first part contains a brief introduction to some of the issues involved in mapping census data: 1) choice of aggregation level and scale, and 2) arbitrary spatial boundaries, MAUP and areal averaging (also see section 1.3.1). The second part describes some common types of maps that are available for mapping ethnic census data, e.g. examples are provided for illustrative purposes only.

4.1.1. Choice of Aggregation Level and Scale

The first step in making a map is to choose the aggregation level and the scale. The aggregation level, however, is dependent upon the scale. Notably, maps are visually tied to scale and aggregation because the eye can only process a limited amount of information. "The choice of how many places to adopt and show on paper does affect the techniques we can use for the visualization simultaneously, and the degree of detail with which we can investigate each one. Basically, the more areas we wish to see at once, the less we can see about each -- there is a finite (if large) amount of detail that a single image can contain and remain comprehensible" (Dorling, 1995, p. 176). With all forms of abstraction, there is a constant struggle to balance the clarity gained by less spatial units, with the detail gained by more spatial units.

> *"Scale relates to the size of the area being studied and determines the level of precision and generalization applied in the study. Microscale studies are done over small earth areas; macroscale studies deal with larger areas. The nature of the inquiry sets the scale, and the scale in turn determines the degree of generalization" (Dent, 1993, p. 11).*

Most segregation studies have been performed on a city or metropolitan scale. However, a metropolitan scale often involves several choices. For example, in Toronto, there are three main boundaries: Metropolitan Toronto, the greater Toronto area (GTA), and Toronto's CMA. Toronto's CMA has been chosen for this study because it is the most natural boundary (see Chapter 2 for definition of CMA).

In segregation studies, the census tract is the most common level of aggregation. Census tracts provide a neighbourhood level of detail and spatial units that are large enough to visualise on a city scale. In addition, the neighbourhood level of aggregation is preferred because most segregation studies are ecologically based. Census subdivisions (municipalities) are also commonly used. Census subdivisions have the advantage of clarity and easily recognized spatial boundaries. As well, CSDs are political boundaries. Federal electoral maps are not commonly used outside electoral studies. Enumeration areas have become more popular because new technology has allowed the use of smaller spatial units. However, some enumeration areas are so small that they are not visible at a city scale. In fact, many are barely visible at very large scales (e.g. 1:30000). Also, the increased number of units at the enumeration level is very time consuming and costly. Therefore, for practical reasons, census tracts still remain the spatial units of choice for most urban mapping. It should be noted, however, that important information can be lost when moving from enumeration areas to census tracts. This is related to the problem of arbitrary census boundaries (see Chapter 1).

4.1.1 Arbitrary Spatial Boundaries, MAUP and Areal Averaging

> "Arbitrary boundaries, however, have their greatest influence on the impression gained in their use for portraying statistics, not in their use for calculating them. If the whole of the country is shaded dark grey because the levels of unemployment are high in one of its towns, is our image accurately reflecting reality" (Dorling, 1995, p. 177).

Figure 4.1 shows the effect of areal averaging by revealing the difference between using census tracts and enumeration areas. Black single-origin data has been used because the spatial distribution of this group consists of small clusters that are scattered throughout the city. In Figure 4.1b, the spatial coverage of the Black group is more pervasive than in Figure 4.1a. In Figure 4.1b, the Black group is visually dominant in the Keele-Finch-Dufferin-Steeles block, while in Figure 4.1a, the Black group is not represented in this area at all. Someone looking at Figure 4.1b would therefore be given a misleading impression of the spatial distribution of the Black group within this area. This loss of information is related to MAUP (modifiable areal unit problem). Unfortunately, spatial patterns are often a product of the way that data is aggregated rather than the underlying spatial distribution. The modifiable areal unit problem (MAUP) further impacts upon thematic mapping because "the greatest emphasis is given to those places containing fewest people -- where the arbitrary movement of boundaries can have the most severe effect on the values calculated" (Dorling, 1995, p. 180). For the most part, however, we are limited to the boundaries established by the census units themselves. Unfortunately, there is little that

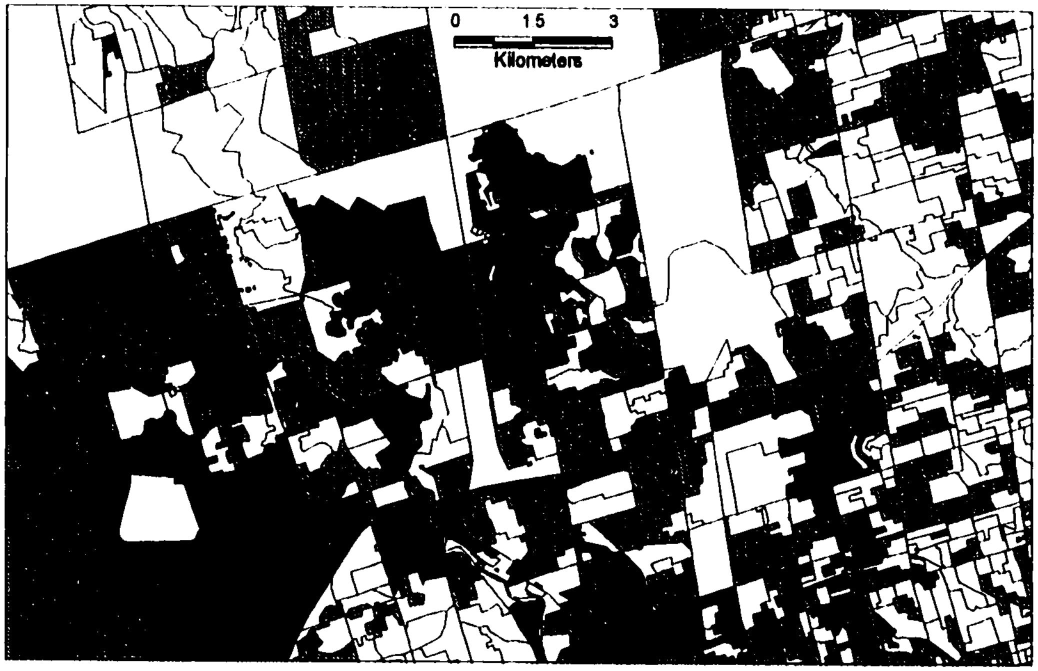
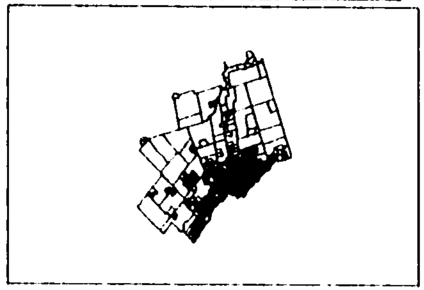


Figure 4.1a): Choropleth Map of People of Single Black Origin (1991) by Enumeration Area.

Black
Single Origin, 1991

- 0.079 to 0.233 (75)
- 0.035 to 0.079 (163)
- 0.001 to 0.035 (462)
- 0 to 0 (112)



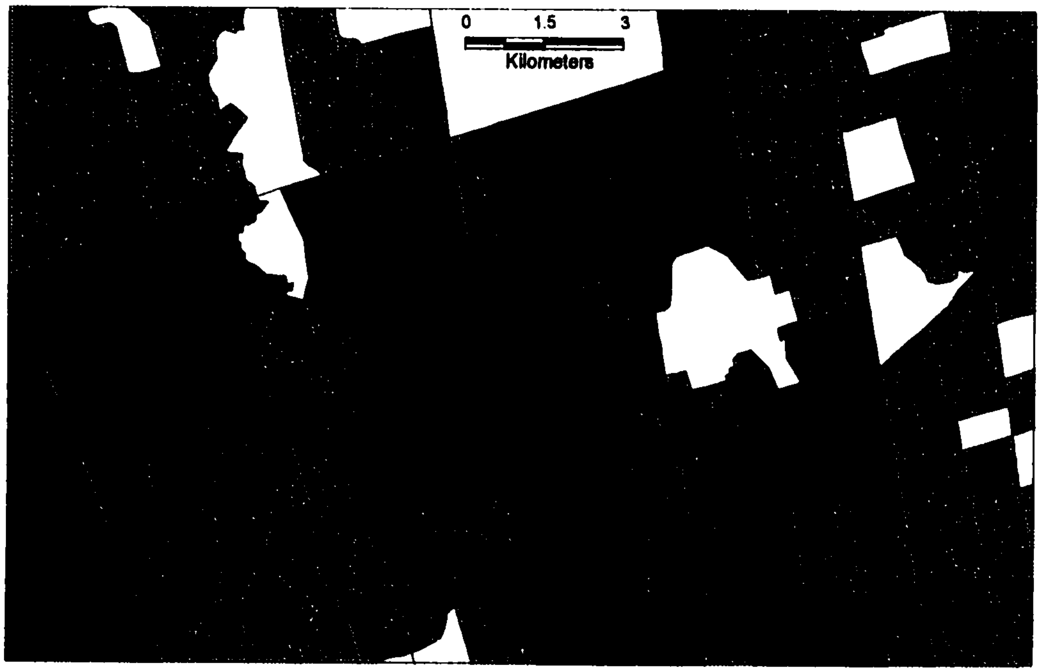


Figure 4.1b): Choropleth Map of People of Single Black Origin (1991) by Census Tract

Black
Single Origin, 1991

- 0.079 to 0.233 (75)
- 0.035 to 0.079 (163)
- 0.001 to 0.035 (462)
- 0 to 0 (112)

Sources:
Statistics Canada, 1991
Wilfrid Laurier University
Compusearch

can be done to rectify the modifiable areal unit problem. However, there are ways of lessening the problem. One way includes producing multiple views of the same data with different boundaries as a way of illustrating the effects of MAUP (Dorling, 1995, p. 170).

4.2. Types of Maps

There are many different types of maps that can be used to represent ethnic census data. The following have been produced in this study to illustrate some of the alternatives available[i].

Univariate Maps

- Areal (Choropleth & Dasymetric)
- Dot Distribution
- Surface Representation
- Proportional Point

Multivariate Maps

- Proportional Pie Chart
- Nominal Areal

Cartograms and Statistical Maps

- Location Quotient
- Cartogram

[i] The maps produced in this study are meant as alternatives and do not represent all the of possibilities available. Furthermore, within each map type, there are critical design issues/choices that can result in numerous versions of the same map.

4.2.1. Univariate Maps

4.2.1.1. Areal (Choropleth & Dasymetric)

The areal representation (or choropleth map) is the most common way of representing census data because it portrays the way that data is collected. Also, it is the most common way of representing population density. However, the choropleth representation is not thought to be the most appropriate model for representing census data (Martin, 1991) [see nature of census boundaries], because the areal representation is not of the same object class as the population (MacEachren, 1994). In addition, areal maps have the problem of placing emphasis upon the largest areal units which are usually rural and unpopulated. A hybrid called the dasymetric choropleth map tries to address the main problem associated with choropleth maps. This type of map attempts to show inter-spatial variation by using residential area only. "Dasymetric mapping means to shade the pieces of land upon which people live" (Dorling, 1995, p. 176). The dasymetric technique is an older technique that was developed in the late 19th century.

> *"Dasymetric mapping, where only the points where people actually live are shaded, have long been advocated to produce more realistic pictures" (Dorling, 1995, p. 177).*

With the choropleth map, data needs to be classified in a clear, meaningful way. Although there are many ways of classifying data, the natural breaks method is usually preferred. The natural breaks method offers the advantage of grouping the phenomena into homogeneous or natural categories. Other methods commonly used are quintals, quartiles, standard deviations, etc. The appropriate classification depends upon the purpose of the study and the phenomena being mapped. However, the quartile and equal interval methods should generally be avoided because they lack the ability to group homogeneous values. Also, when classifying data, ratio data should be used instead of absolute values and the colour scheme should show a hierarchy. This is usually achieved by changing the colour saturation. Overall, the areal classification has the advantage of identifying areas that have a higher proportion of ethnic groups. However, when data is classified into categories, it is more open to manipulation.

Figure 4.2.1: A comparison of a) Toronto's CMA total area to b) the residential area within Toronto's CMA

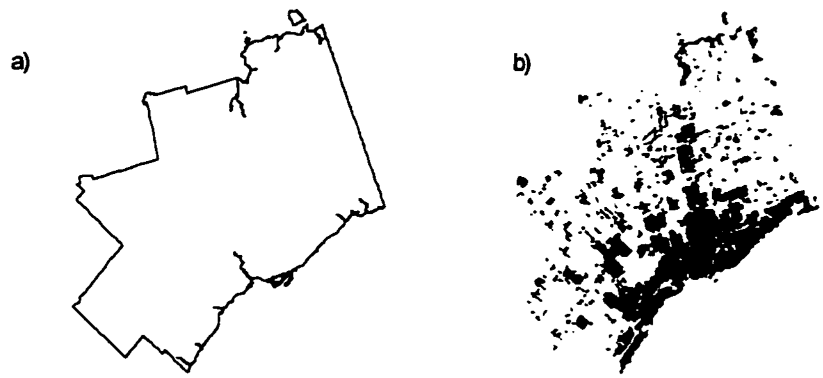


Figure 4.2.2: A comparison of a) regular enumeration units with b) dasymetric (residential land only) enumeration units

The areal, choropleth maps that were produced in this study include normal choropleth and dasymetric versions. Figure 4.2 shows the residential area differentiation that was used to make the dasymetric maps: Figure 4.2.1a shows the whole study area; Figure 4.2.1b shows the residential area only (see Chapter 2); Figure 4.2.2a shows the original enumeration census boundaries; and Figure 4.2.2b shows the new dasymetric census boundaries. Intersecting the original census boundaries with the residential areas produced the new residential boundaries. There is a dramatic difference visually between the old census units and the new dasymetric units, especially in the rural areas. Unfortunately at this scale, it is difficult to see the delineation between areas within Metropolitan Toronto. The distinction becomes clearer in maps displayed at smaller scales.

Figure 4.3 shows two choropleth maps of the Chinese group (normal and dasymetric). There is a dramatic difference visually between the rural areas in the two maps. Although Figure 4.3a shows that most of the population of this group is located outside Metropolitan Toronto, in reality only 26.3% is located outside. In contrast, this area is not as visually dominant in Figure 4.3b. The dasymetric map would have been a better representation in this case because it does not let large unpopulated land dominate. However, dasymetric maps are more time consuming because they require further investigation into the population distribution of the study area. Also, although the dasymetric map is an improvement over the traditional choropleth map, the areas with the least population are still visually dominant on the map.

> *"Unfortunately, as we more accurately portray physical reality, we are left with less and less space in which to show social reality. Even in the most densely populated part of Britain, what dasymetric mapping shows best are those areas which it leaves colourless. The parks, rivers and roads of the capital stand out most clearly on a dasymetric map of any variable for a city like London, and then most prominent are the isolated places, which usually contain least people and for which averages and proportions are least meaningful" (Dorling, 1995, p. 180).*

Figure 4.3a: Standard Choropleth Map of People of Single Chinese Origin (1991), in Toronto's CMA

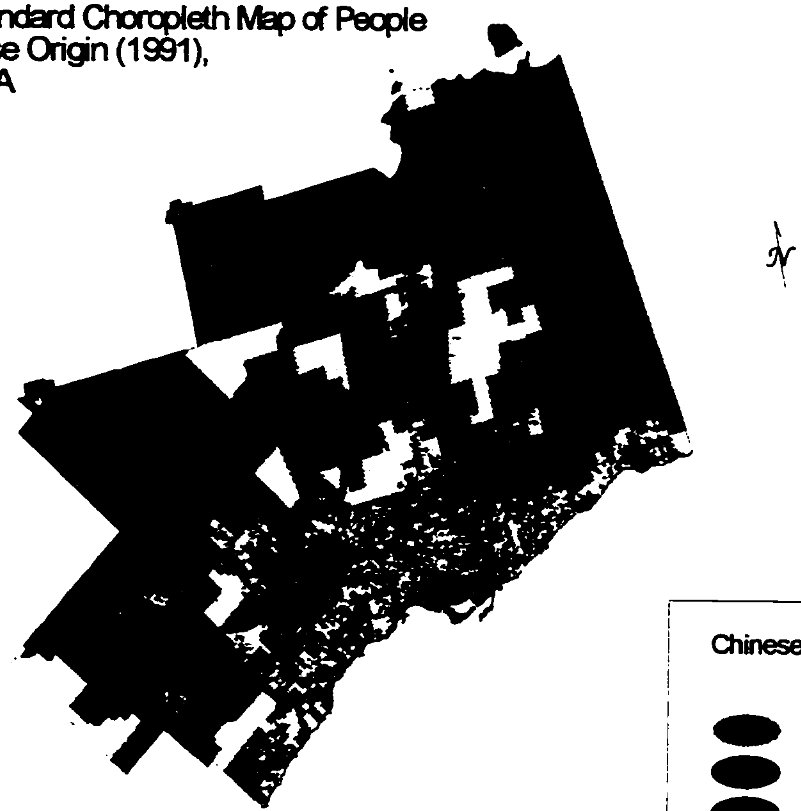
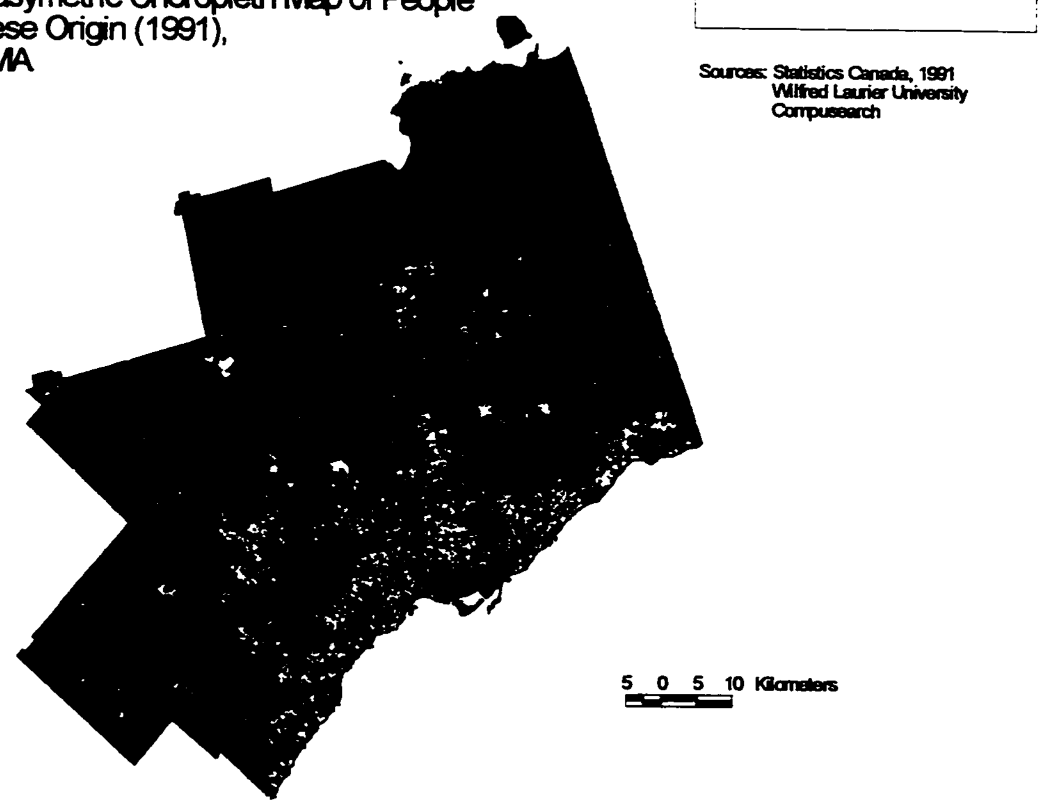


Figure 4.3b: Dasymetric Choropleth Map of People of Single Chinese Origin (1991), in Toronto's CMA



Notably, in both maps in Figure 4.3, the small enumeration units (e.g. apartments buildings) are barely visible or not visible at the scale used.

4.2.1.2. Dot Distribution

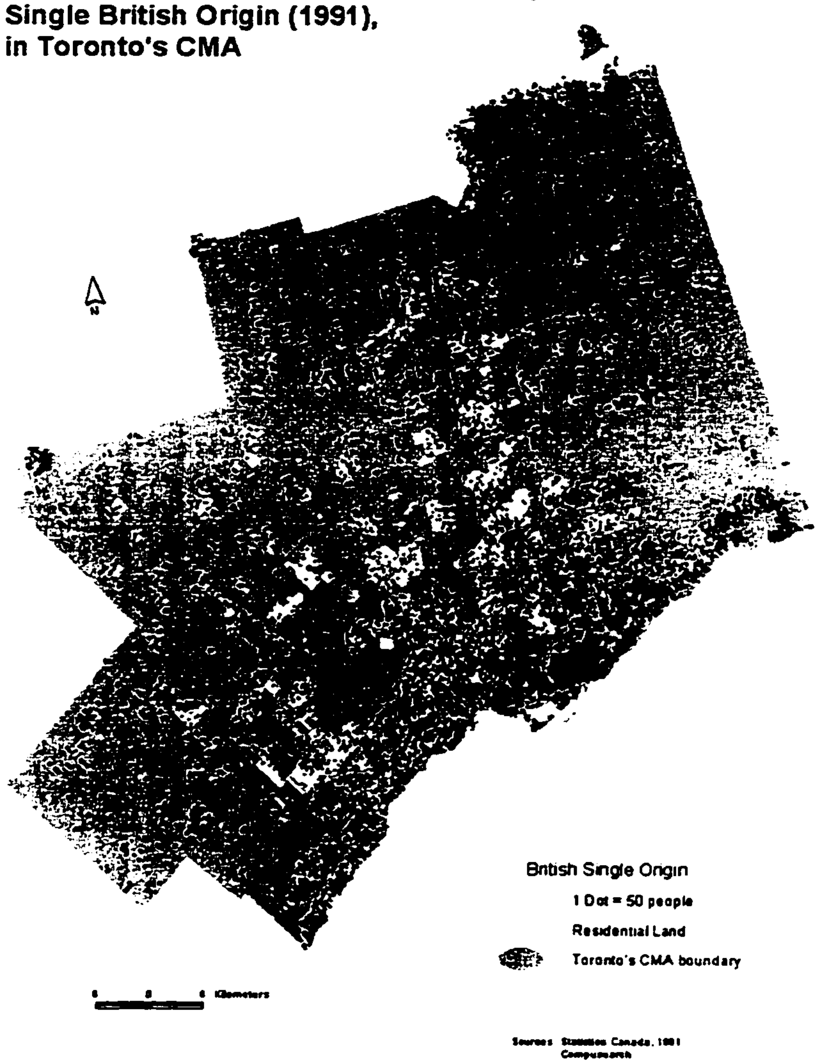
The dot distribution map is not as common as the standard choropleth map. However, the dot distribution map has the advantage of visualizing population density, being of the same object class as the data, and not giving visual emphasis to unpopulated areas. Despite advantages, however, dot distribution maps cannot identify areas that have a higher proportion of ethnics. This is usually accomplished by using a choropleth map. Also, dot distribution maps can be very time consuming because they require finding populated areas within the spatial units. In addition, the placement as well as the size and value of the dots can dramatically change the visual impact of the dot distribution map (see Dent, 1993). With dot distribution maps, it is preferred that (a) small spatial units are used and (b) dots are randomly distributed within residential areas only. Interpretation errors can occur when dots are placed in unpopulated areas. For example, a dot distribution map that randomly distributed dots over all of California would give a misleading impression when it placed dots/population over the Rocky Mountains. Unfortunately, most standard mapping packages do not allow users to place dots manually. Instead, the dots are placed randomly within the spatial units by the software used. To overcome this, the dot distribution maps in this study were produced from the dasymetric enumeration area polygons. The dots were generated randomly within the dasymetric census polygons by using standard mapping software (e.g. MapInfo and ArcView).

Dot distribution maps have been produced for all ethnic groups within this study because these maps are considered the best for comparative purposes (see Figures 4.4 through 4.17). The same dot representation has been used for all the maps, e.g. one dot = 50 people. Notably, ethnic groups have been placed into groups of low, intermediate and high segregation in order to visualize the gradual increase in segregation.[ii] Index values are also included beside each dot distribution map. The only exception is the British group, because this group has been used as the comparison group y.

[ii] Maps are not all in the same segregation order as listed in Chapter 3.

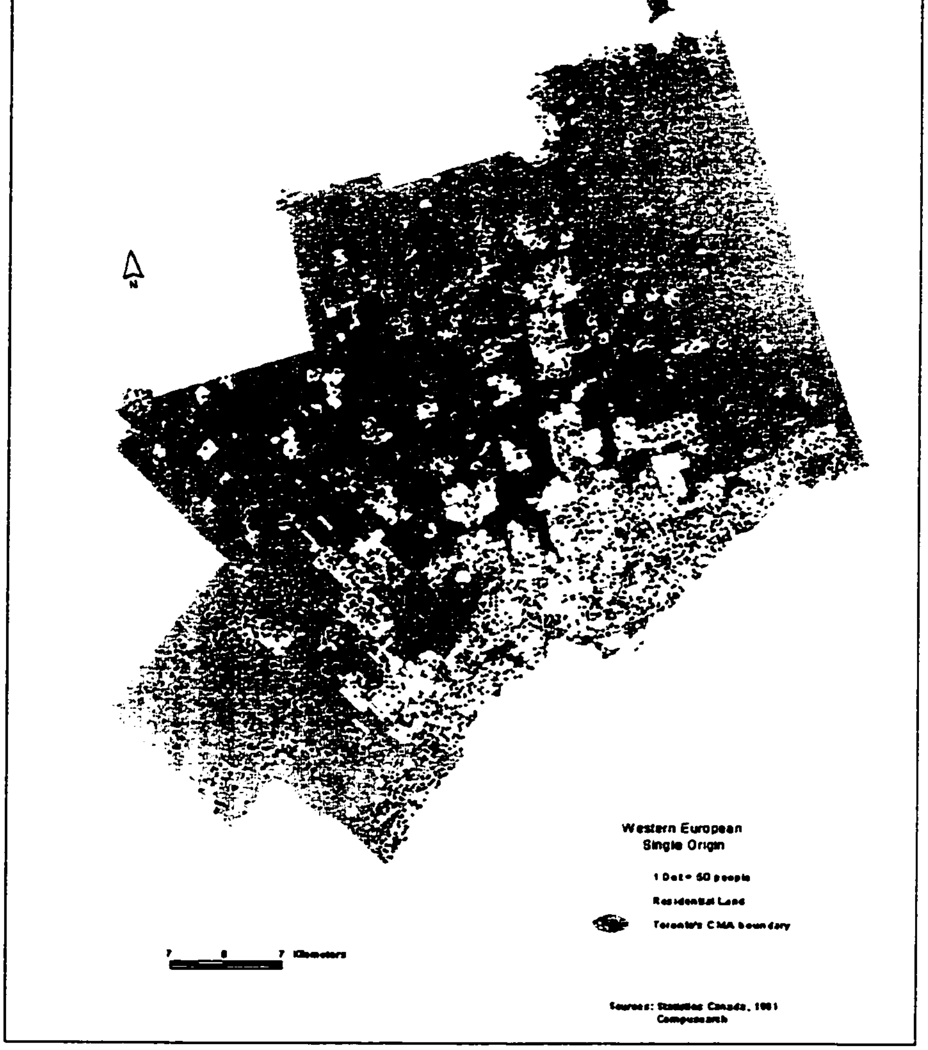
Ethnic Groups that reveal low levels of segregation

Figure 4.4: Dot Distribution Map of People of Single British Origin (1991), in Toronto's CMA



The spatial distribution of the **British** group (Figure 4.4) is comprehensive and even. Although this group is strongly represented throughout the study area, it has less representation in the western and eastern areas of the city (Metropolitan Toronto). The British is therefore one of the least segregated groups within Toronto's CMA. The British and/or English group has been used as a benchmark (e.g. used as comparison group y).

Figure 4.5: Dot Distribution Map of People of Single Western European Origin (1991), in Toronto's CMA



Western European

index	EA
d-index	0.395
xP y	0.161
RCO	-.446
ACE	0.56
SP	0.144

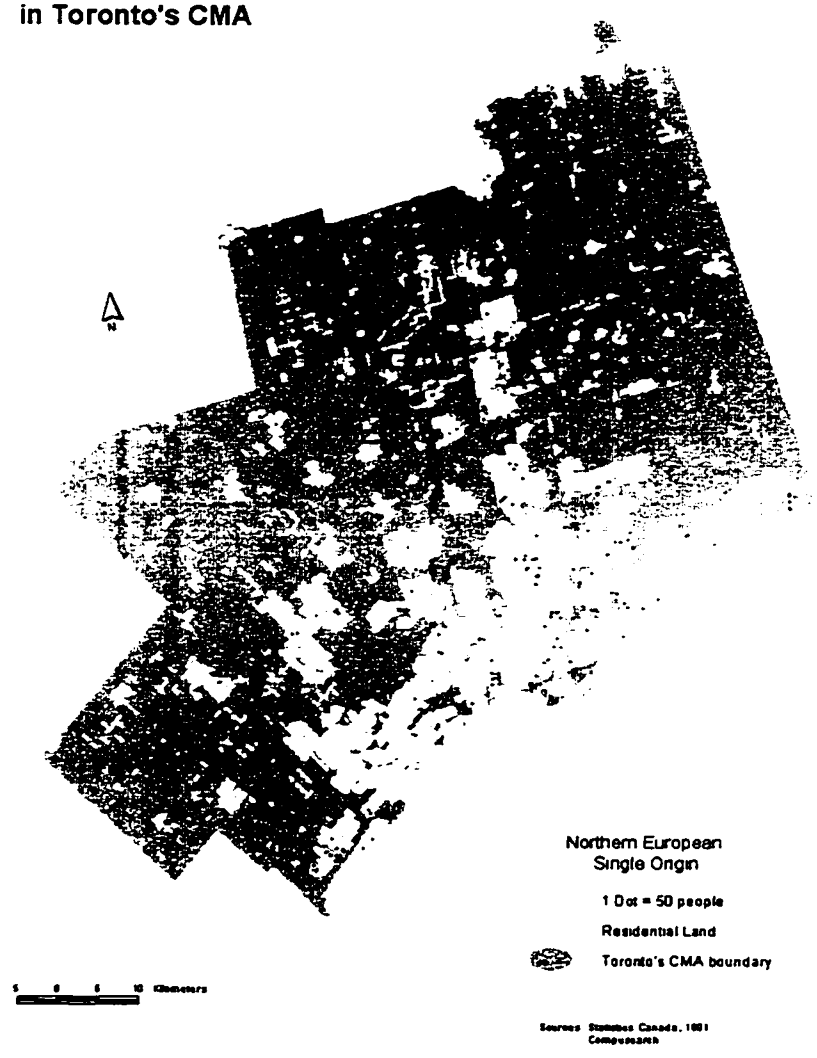
index	CT
d-index	0.252
xP y	0.159
RCO	-.357
ACE	0.56
SP	0.168

index	FED
d-index	0.153
xP y	0.149
RCO	-.160
ACE	0.54

index	CSD
d-index	0.151
xP y	0.148
RCO	-.04
ACE	0.63

The **Western European** group (Figure 4.5) reveals a spatial pattern that is similar to the British. That is to say that the spatial distribution of this group is even and dispersed. The Western European group is the only ethnic group that has most of its population located outside Metropolitan Toronto. This accounts for the low score that they received on the centralization dimension.

Figure 4.6: Dot Distribution Map of People of Single Northern European Origin (1991), in Toronto's CMA



Northern European

index	EA
d-index	0.786
xPy	0.159
RCO	0.082
ACE	0.66
SP	0.124

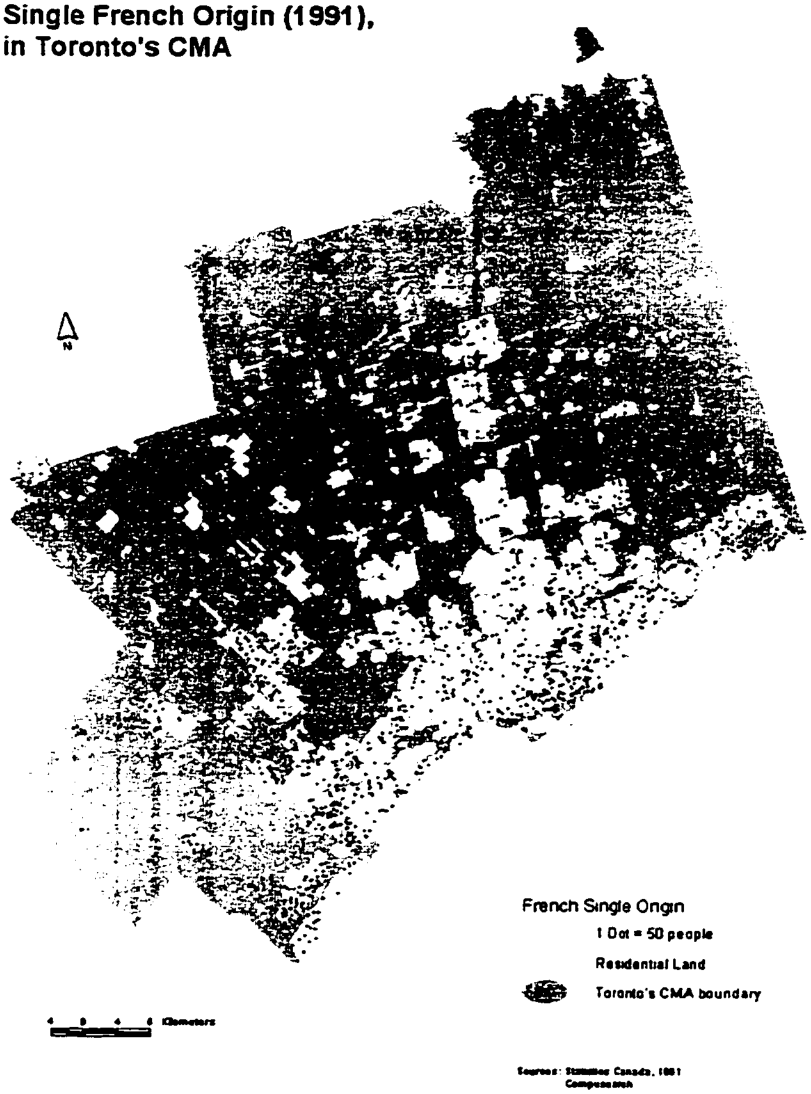
index	CT
d-index	0.423
xPy	0.152
RCO	0.118
ACE	0.66
SP	0.145

index	FED
d-index	0.170
xPy	0.140
RCO	0.11
ACE	0.63

index	CSD
d-index	0.12
xPy	0.139
RCO	0.09
ACE	0.63

The **Northern European** map (Figure 4.6) illustrates the small population of this group. Although the spatial distribution of this group is fairly even, it is limited to central Toronto, the Lakeshore, and the Yonge Street corridor. Notably, there are many areas within the city where this group is not represented.

Figure 4.7: Dot Distribution Map of People of Single French Origin (1991), in Toronto's CMA



French

index	EA
d-index	0.452
xPy	0.15
RCO	0.472
ACE	0.68
SP	0.135

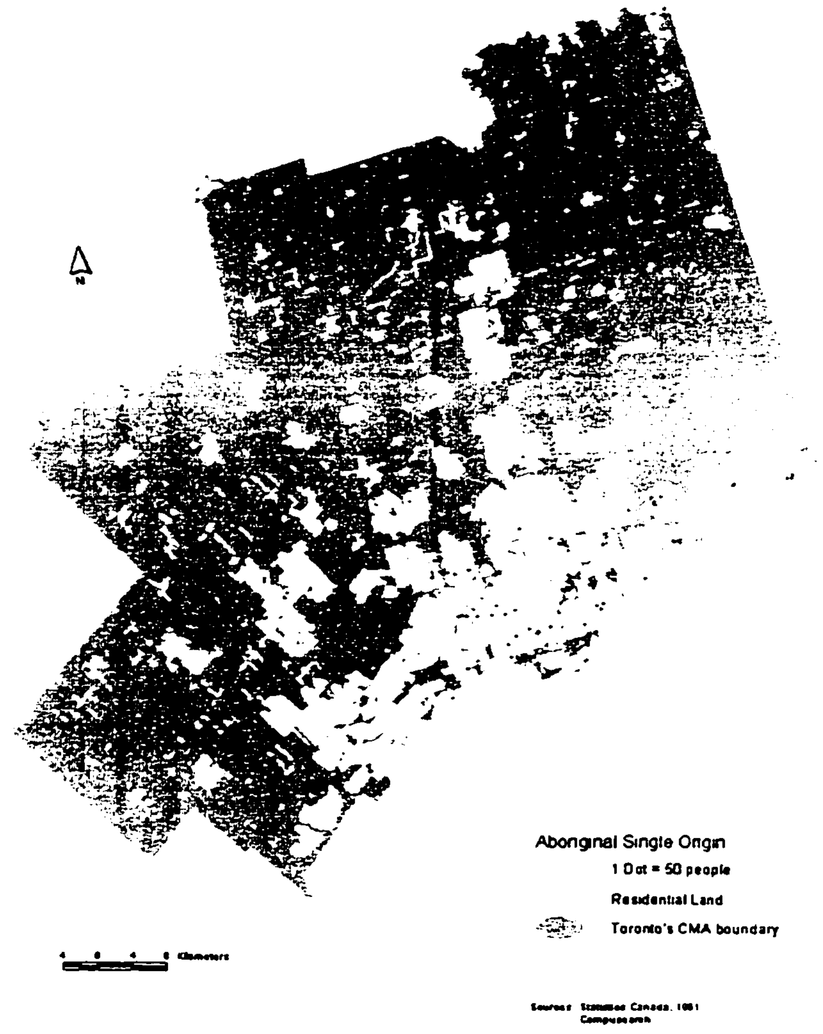
index	CT
d-index	0.240
xPy	0.146
RCO	0.433
ACE	0.68
SP	0.155

index	FED
d-index	0.127
xPy	0.140
RCO	0.260
ACE	0.66

index	CSD
d-index	0.113
xPy	0.138
RCO	0.050
ACE	0.66

The **French** map (Figure 4.7) shows that the French group has a small population within Toronto's CMA. The limited population of this group is randomly dispersed throughout the study area.

Figure 4.8: Dot Distribution Map of People of Single Aboriginal Origin (1991), in Toronto's CMA



Aboriginal

index	EA
d-index	0.912
xPy	0.141
RCO	0.467
ACE	0.0
SP	0.123

index	CT
d-index	0.595
xPy	0.138
RCO	0.530
ACE	0.0
SP	0.144

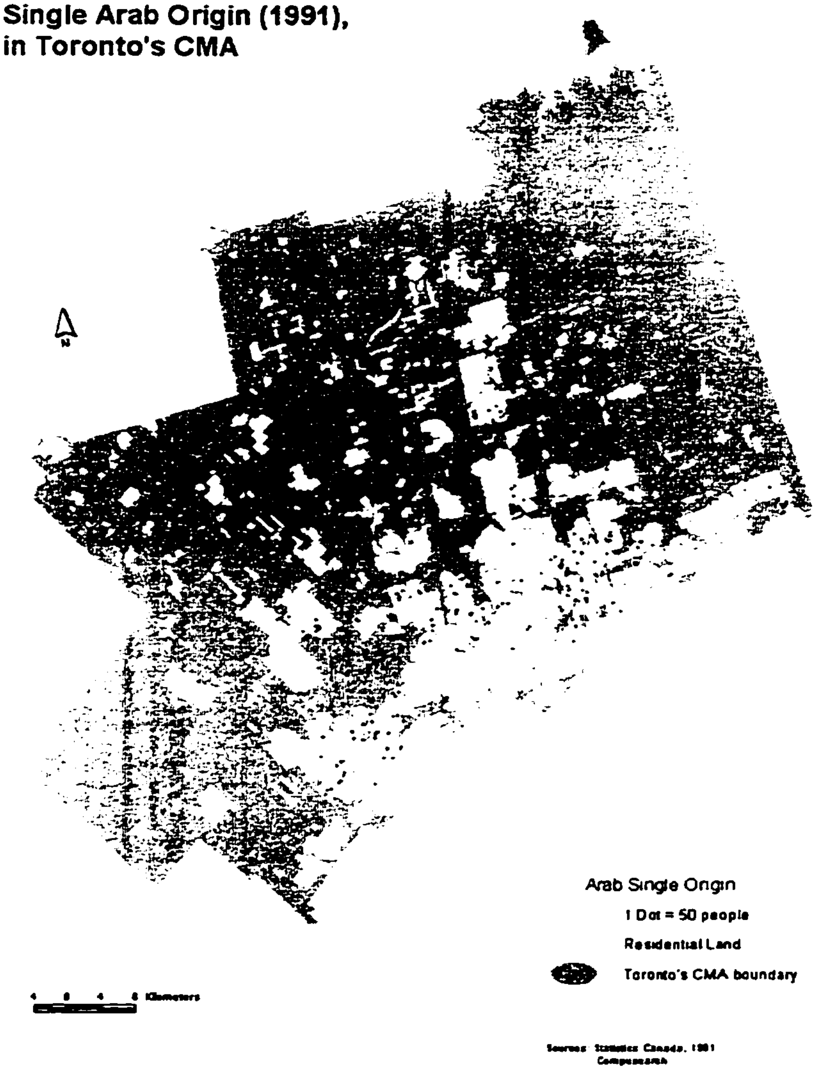
index	FED
d-index	0.205
xPy	0.135
RCO	0.340
ACE	0.0

index	CSD
d-index	0.170
xPy	0.132
RCO	0.140
ACE	0.0

The **Aboriginal** population (Figure 4.8) is so small that it is hard to establish any spatial pattern. However, despite limited coverage, the spatial distribution of this group appears to be even and unclustered.

Ethnic groups that reveal intermediate levels of segregation

Figure 4.9: Dot Distribution Map of People of Single Arab Origin (1991), in Toronto's CMA



Arab

index	EA
d-index	0.901
xP y	0.119
RCO	0759
ACE	0.78
SP	0.126

index	CT
d-index	0.623
xP y	0.124
RCO	0.693
ACE	0.78
SP	0.147

index	FED
d-index	0.335
xP y	0.119
RCO	0.610
ACE	0.77

index	CSD
d-index	0.300
xP y	0.122
RCO	-0.01
ACE	0.77

The **Arab** (Lebanese) map (Figure 4.9) shows that the spatial distribution of this group is very limited. Specifically, spatial coverage appears to be restricted to the northern areas of Metropolitan Toronto. Small clusters appear towards the west along the city border (Steeles Avenue West).

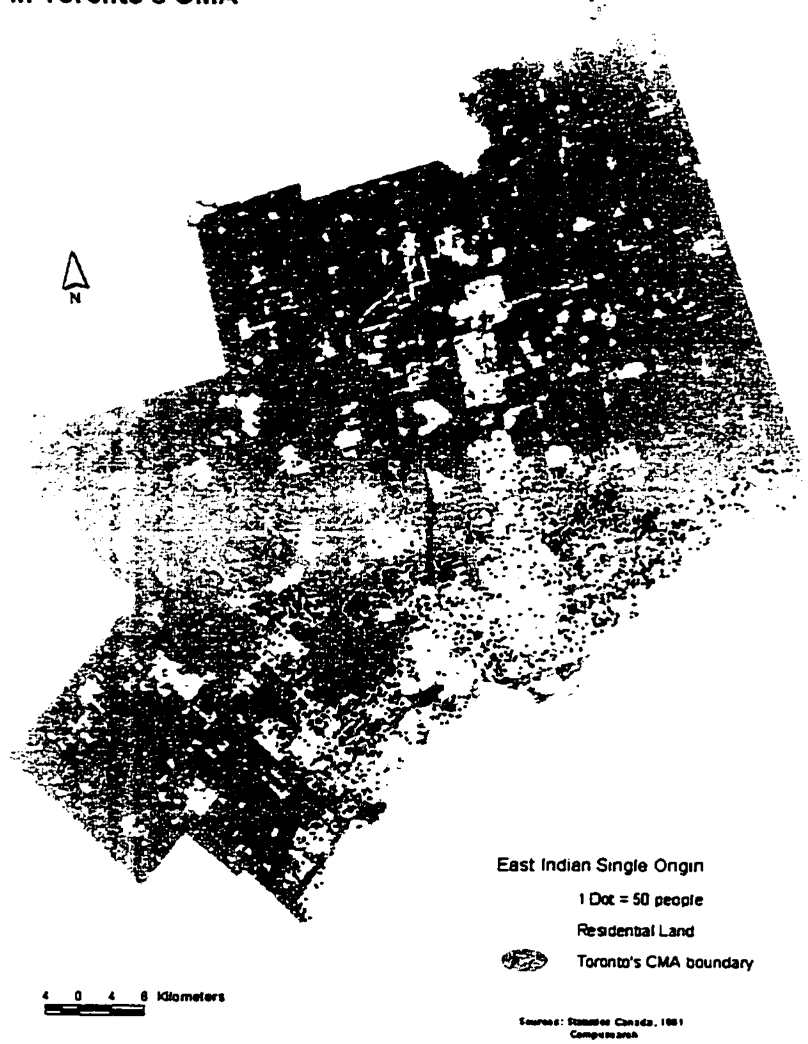


Figure 4.10: Dot Distribution Map of People of Single East Indian Origin (1991), in Toronto's CMA

East Indian

index	EA
d-index	0.473
xP y	0.127
RCO	0.549
ACE	0.78
SP	0.18

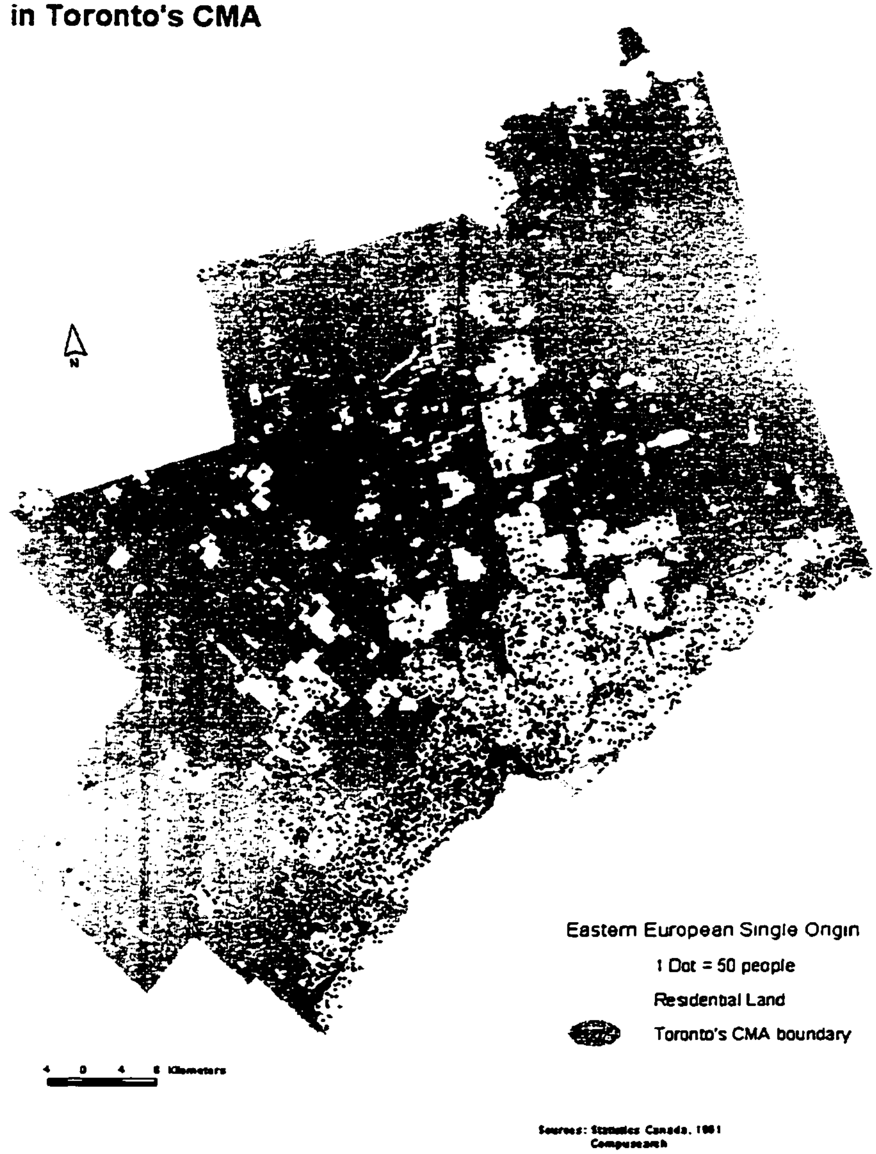
index	CT
d-index	0.365
xP y	0.103
RCO	0.496
ACE	0.78
SP	0.208

index	FED
d-index	0.299
xP y	0.124
RCO	0.53
ACE	0.75

index	CSD
d-index	0.25
xP y	0.127
RCO	0.1
ACE	0.75

The **East Indian** map (Figure 4.10) shows that this group has an intermediate level of segregation. Although the spatial distribution of this group is fairly scattered, there are distinct clusters in Brampton and the Steeles-Kipling area. Mississauga, Markham, and Scarborough also have a large number of East Indians. The East Indian population has limited representation in rural areas.

Figure 4.11: Dot Distribution Map of People of Single Eastern European Origin (1991), in Toronto's CMA



Eastern European

index	EA
d-index	0.473
xPy	0.127
RCO	0.549
ACE	0.74
SP	0.180

index	CT
d-index	0.365
xPy	0.13
RCO	0.496
ACE	0.74
SP	0.208

index	FED
d-index	0.299
xPy	0.129
RCO	0.530
ACE	0.72

index	CSD
d-index	0.25
xPy	0.127
RCO	0.1
ACE	0.72

The **Eastern European** map (Figure 4.11) shows that most of this group is located in Etobicoke and Mississauga. Although the coverage of this group is fairly even, a large cluster appears in Southern Etobicoke.

Figure 4.12: Dot Distribution Map of People of Single Black Origin (1991), in Toronto's CMA

Black

index	EA
d-index	0.571
xP y	0.102
RCO	0.810
ACE	0.78
SP	0.167

index	CT
d-index	0.425
xP y	0.107
RCO	0.716
ACE	0.78
SP	0.202

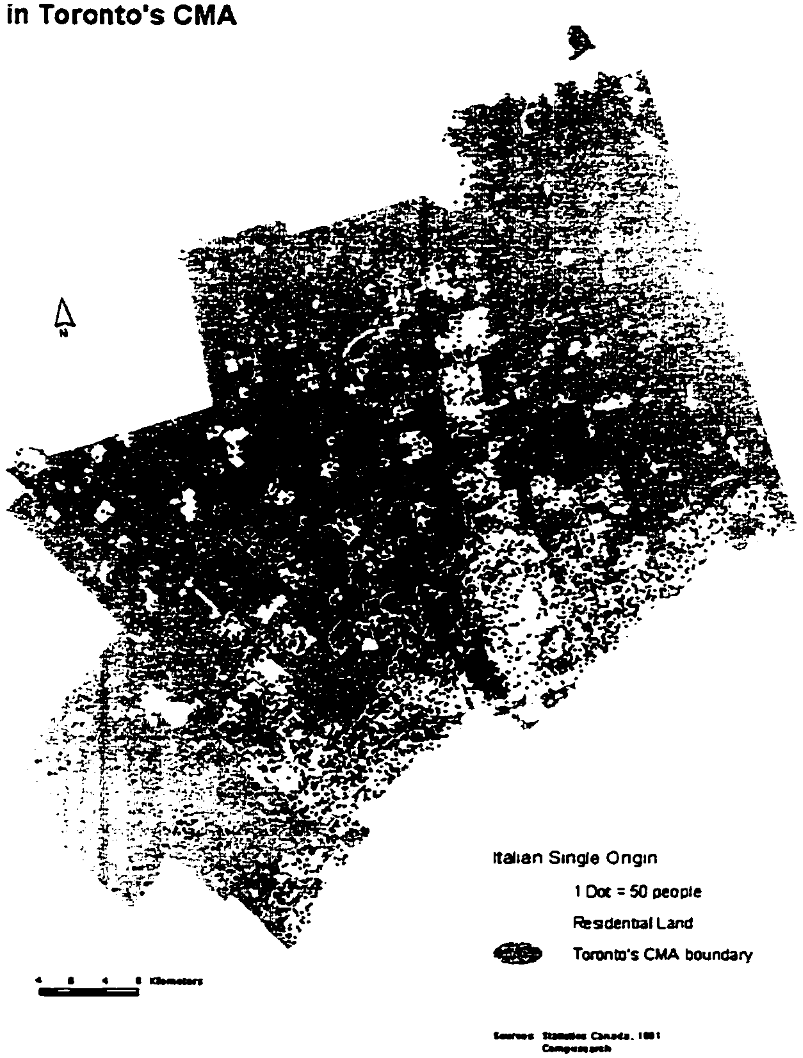
index	FED
d-index	0.283
xP y	0.119
RCO	0.610
ACE	0.77

index	CSD
d-index	0.214
xP y	0.123
RCO	0.050
ACE	0.77

Figure 4.12 shows that most of the population of the **Black** group is located within Metropolitan Toronto, Mississauga, Brampton, South Pickering and Ajax. Stronger representation is revealed in Downsview and Scarborough (small clusters are scattered throughout these areas). Notably, there is virtually no representation in the rural and northern areas of the study. Also, there appears to be little representation in the north-central (uptown) area of Metropolitan Toronto.

Ethnic groups that reveal high levels of segregation

Figure 4.13: Dot Distribution Map of People of Single Italian Origin (1991), in Toronto's CMA



Italian

index	EA
d-index	0.568
xP y	0.09
RCO	0.272
ACE	0.76
SP	0.236

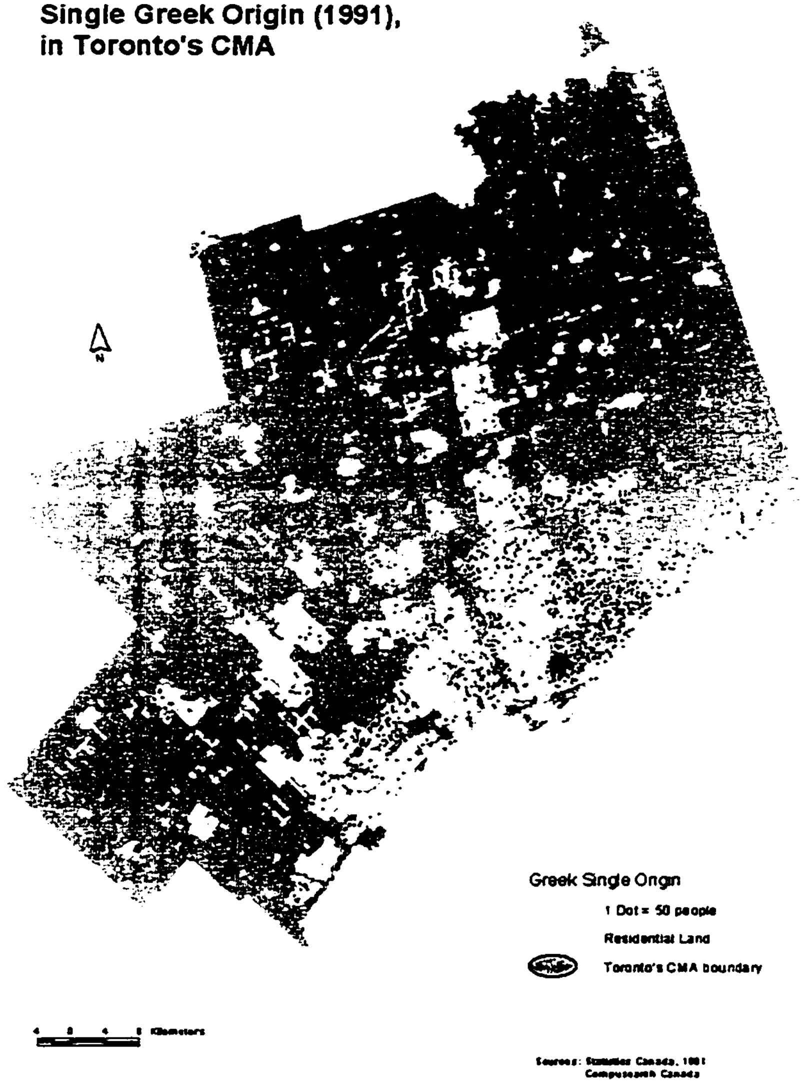
index	CT
d-index	0.495
xP y	0.096
RCO	0.353
ACE	0.76
SP	0.327

index	FED
d-index	0.403
xP y	0.111
RCO	0.14
ACE	0.74

index	CSD
d-index	0.312
xP y	0.111
RCO	-0.02
ACE	0.74

The **Italian** map (Figure 4.13) illustrates the strong coverage of this group within the study area. Although the spatial distribution of this group is mainly concentrated in the western areas of the city (York, Downsview), it extends as far northwest as Woodbridge. There appears to be very little representation within the (former) city of Toronto. This group is distinct in that it is one of the few ethnic groups that has coverage in the rural north.

Figure 4.14: Dot Distribution Map of People of Single Greek Origin (1991), in Toronto's CMA



Greek

index	EA
d-index	0.618
xPy	0.111
RCO	0.682
ACE	0.83
SP	0.137

index	CT
d-index	0.442
xPy	0.116
RCO	0.679
ACE	0.83
SP	0.158

index	FED
d-index	0.302
xPy	0.124
RCO	0.55
ACE	0.81

index	CSD
d-index	0.242
xPy	0.124
RCO	0.2
ACE	0.81

The **Greek** group (Figure 4.14) has visible clusters in Metropolitan Toronto and the east. A large cluster is revealed along the Danforth. Although there is limited representation outside Metropolitan Toronto, there is virtually no representation in the rural/peripheral regions.

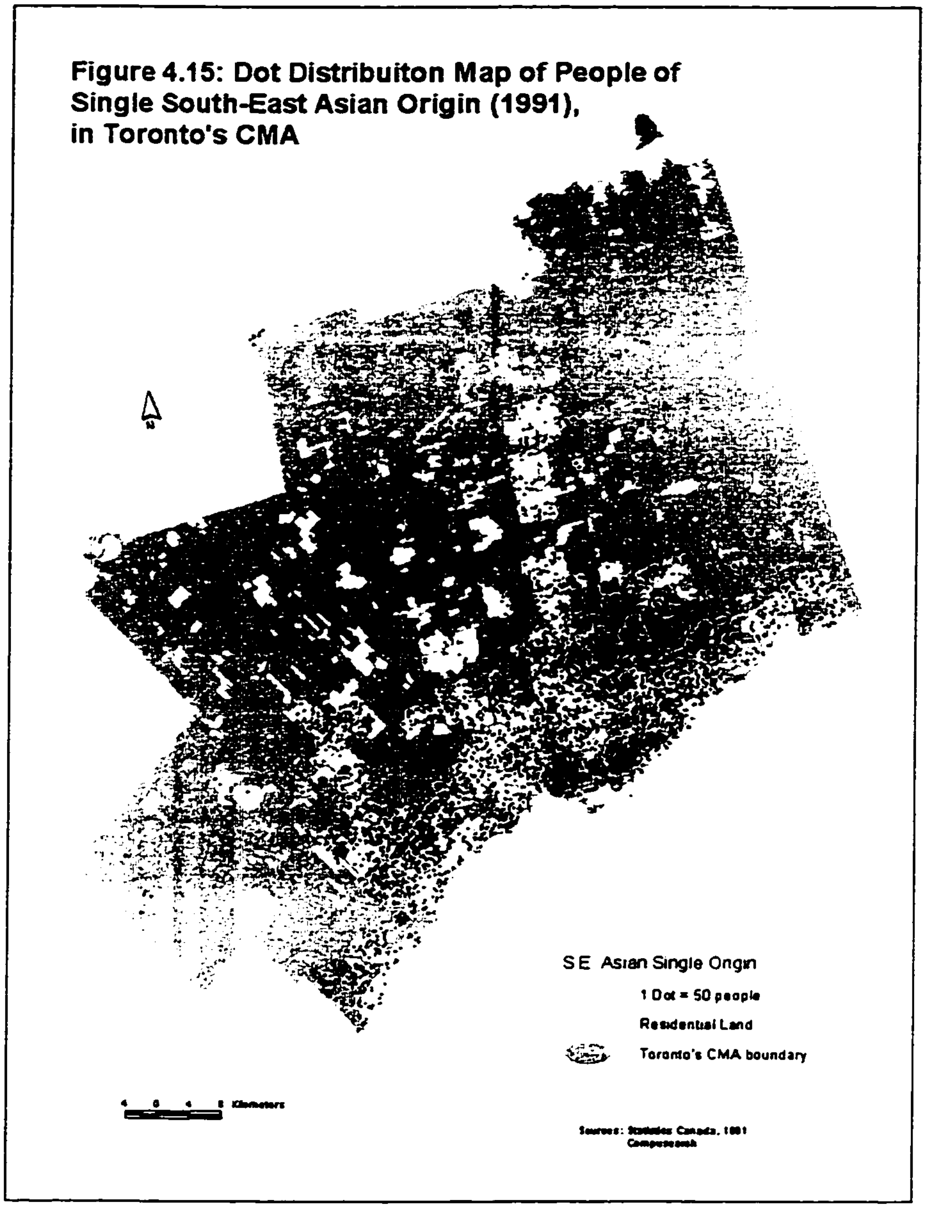


Figure 4.15: Dot Distribuiton Map of People of Single South-East Asian Origin (1991), in Toronto's CMA

S.E. Asian

index	EA
d-index	0.656
xP y	0.088
RCO	0.785
ACE	0.80
SP	0.227

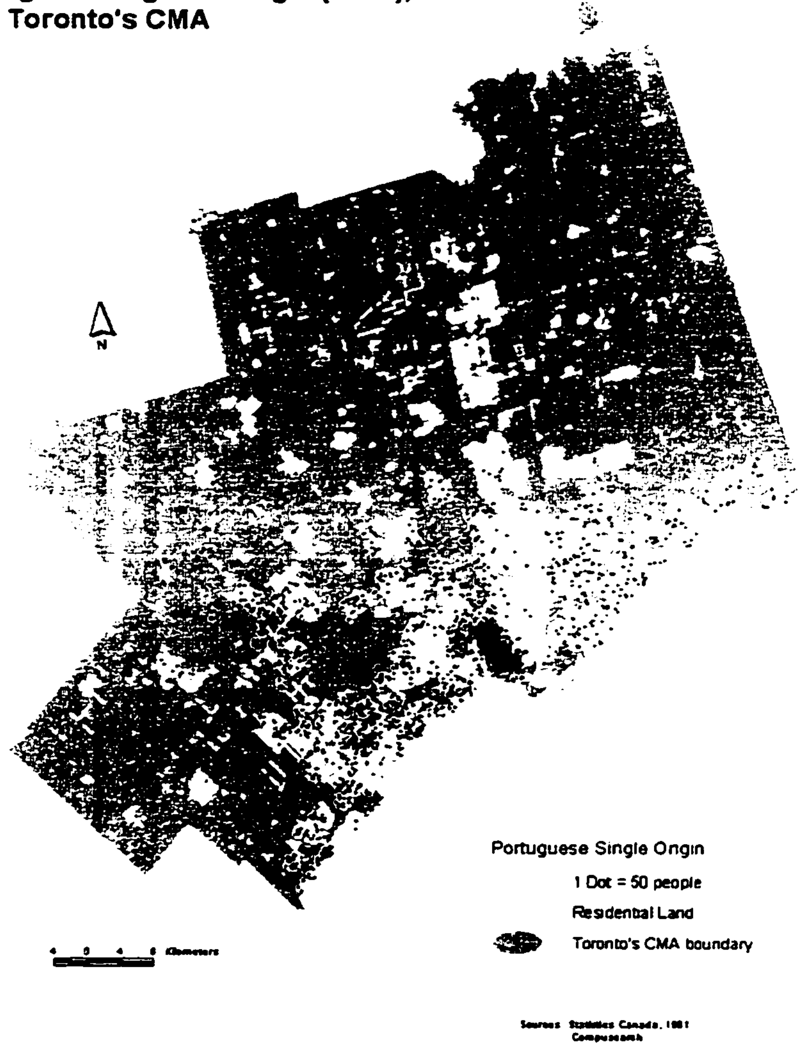
index	CT
d-index	0.577
xP y	0.093
RCO	0.726
ACE	0.8
SP	0.275

index	FED
d-index	0.474
xP y	0.104
RCO	0.61
ACE	0.78

index	CSD
d-index	0.414
xP y	0.124
RCO	0.17
ACE	0.78

The **South East Asian** map (Figure 4.15) reveals a large population within Toronto's CMA. Most of the population of this group is concentrated in Metropolitan Toronto, Mississauga, Brampton, Richmond Hill and Markham. Large clusters appear in Markham and in the (former) city of Toronto. Although this group has large clusters, it also has a significant population that is scattered elsewhere. Therefore, this group is not as segregated as the Portuguese and the Jewish.

Figure 4.16: Dot Distribution Map of People of Single Portuguese Origin (1991), in Toronto's CMA



Portuguese

index	EA
d-index	0.656
xP y	0.088
RCO	0.785
ACE	0.80
SP	0.227

index	CT
d-index	0.577
xP y	0.093
RCO	0.726
ACE	0.8
SP	0.275

index	FED
d-index	0.474
xP y	0.104
RCO	0.61
ACE	0.78

index	CSD
d-index	0.414
xP y	0.124
RCO	0.17
ACE	0.78

The **Portuguese** map (Figure 4.16) shows why this group has the second highest segregation ranking. Most of the population of this group appears in Western Toronto and York (a large cluster) and Bradford (a smaller cluster). In addition, this group has a thin, scattered population in Mississauga, Brampton, Etobicoke and Oakville.

Figure 4.17: Dot Distribution Map of People of Single Jewish Origin (1991), in Toronto's CMA

Jewish

index	EA
d-index	0.811
xP y	0.072
RCO	0.703
ACE	0.87
SP	0.208

index	CT
d-index	0.775
xP y	0.078
RCO	0.756
ACE	0.86
SP	0.389

index	FED
d-index	0.669
xP y	0.109
RCO	0.25
ACE	0.84

index	CSD
d-index	0.548
xP y	0.095
RCO	0.08
ACE	0.84

The **Jewish** map (Figure 4.17) shows why this group is the most segregated group within Toronto's CMA. The Jewish group is mainly clustered along the Bathurst street corridor and north along Bayview Avenue. However, there is little representation elsewhere. The spatial distribution of the Jewish group contrasts dramatically with the spatial distribution of the British.

4.2.1.3. Surface Representation

"Much more useful is [GIS's] ability to re-engineer census geographies in non-straightforward ways, e.g. create zoning systems that are composed of areal entities that are similar in size or shape or social heterogeneity. This is a basic prerequisite for sensible geographical analysis and is one of the greatest gifts that GIS can give the census user: a freedom from the tyranny of fixed arbitrary census geography" (Openshaw, 1995, p. 133).

With new computer technology, data can be manipulated more easily into different types of representation. Because of the current ease in generating surfaces, surface representations are used increasingly to represent census data. However, when census data is represented as a surface, it is necessary to assume that the data is continuous.

"The density or concentration of population can be represented by a three dimensional surface if we agree that the social and economic characteristics of an area are influenced by its neighbouring areas. In other words, the spatial changes are continuous" (Charlton, et al., 1995, 157).

Surface representations come in several formats, including raster (grid), lattice (point) and vector (TIN or contour). However, for clarity, this study will focus on the raster and TIN (Triangulated Irregular Network) formats. The raster representation offers several advantages: (1) that the size of areas is uniform; (2) that large, unpopulated areas are often visually dominant; and (3) that historical changes in political boundaries disrupt the continuity of statistical comparison (Tufte, 1990). Also, the raster data structure is not as precise as the vector data structure. The TIN vector surface model "is a terrain model that uses a sheet of continuous, connected triangular facets based on a Delaunay triangulation of irregularly spaced nodes or observation points. Unlike the altitude matrices, the TIN allows extra information to be gathered in areas of complex relief without the need for huge amounts of redundant data to be gathered from areas of simple relief" (Burrough, 1986, p. 42). Therefore, the data storage requirements for TINs are generally much less than for regular grid surfaces. However, both are easily produced in many standard applications (e.g. ArcInfo, Surfer). Other surface generation techniques are more computer intensive because they involve moving average windows (Langford & Unwin, 1994; Goodchild, et al, 1993; and Bracken and Martin, 1995). Further considerations are the method of interpolation and whether to convert from points to polygons. There are several methods of interpolation (see Burrough, 1986). In this study, two examples have been produced to illustrate the visual impact of surface representations (Figures 4.18 and 4.19).

Figure 4.18 shows a TIN mesh diagram of the Jewish group, while Figure 4.19 shows a grid surface map of the Jewish group. Both were interpolated from the same point file (census tract centroid).

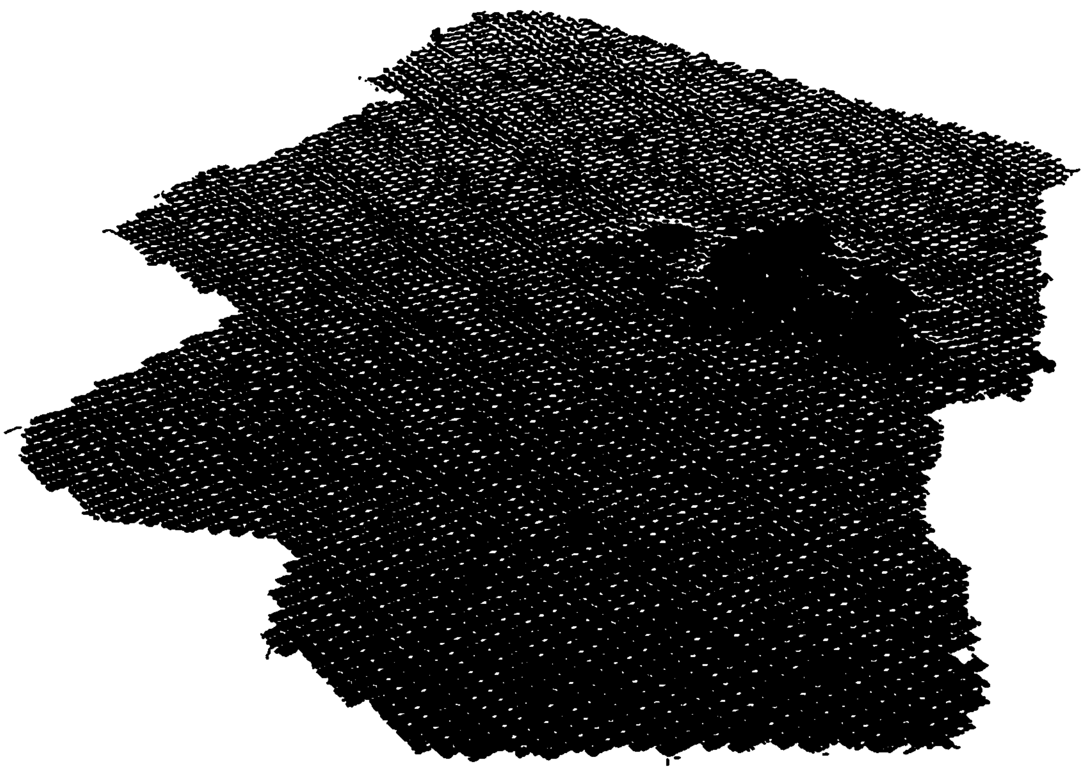


Figure 4.18: Tin mesh map of the Jewish group (single origin, 1991)

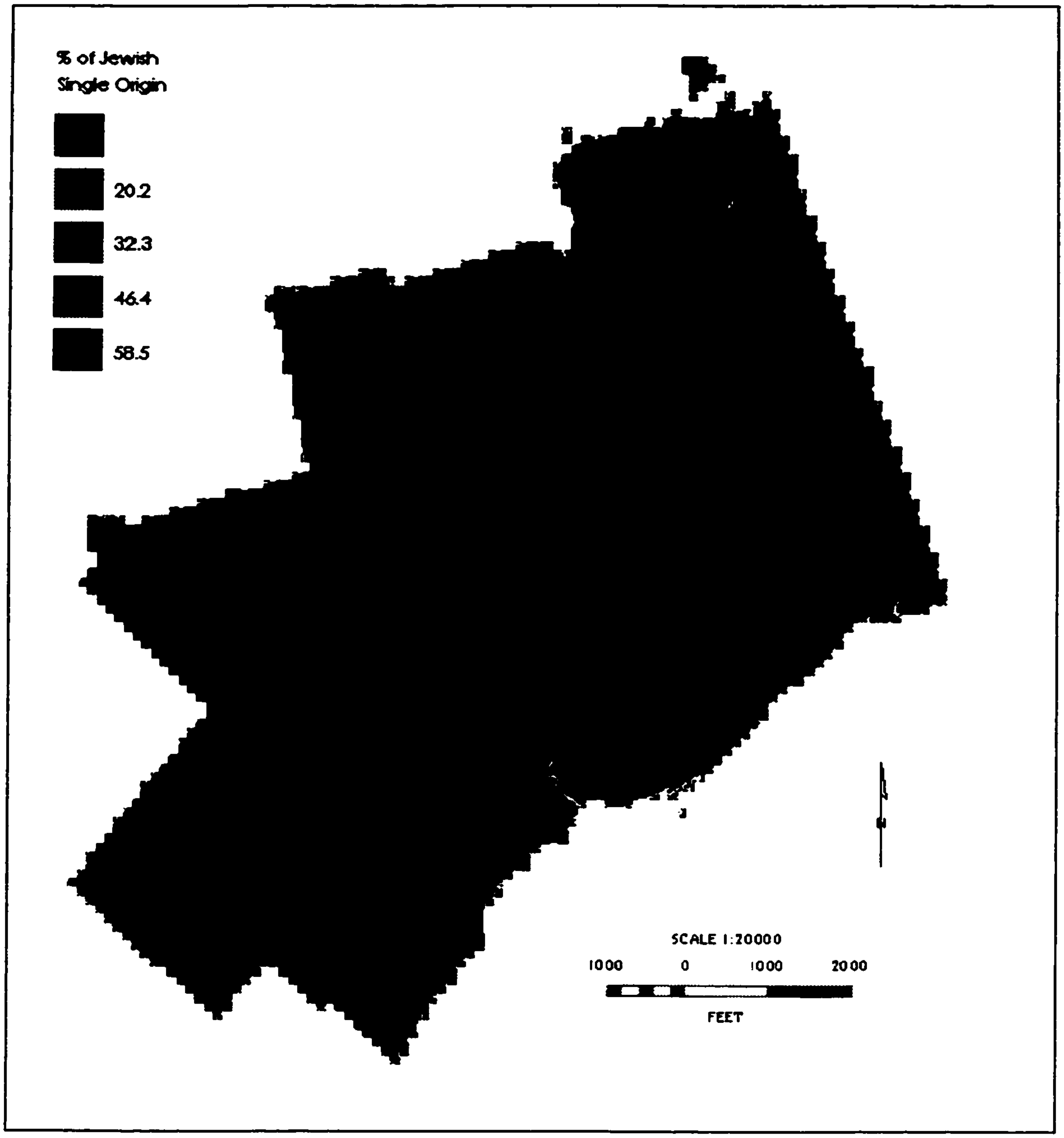


Figure 4.19: Grid Map of People of Jewish Single Origin

Although surface representations have certain advantages, they also have disadvantages. When transforming to a surface model, data is further removed from its origins and therefore more open to manipulation, misrepresentation, and error. This is further accentuated by the fact that different ways of interpolating data can produce varying results. Another disadvantage is that a surface representation shows a continuous distribution. Showing census data as continuous can be somewhat misleading. Although it can be argued that the continuous nature of population is a function of scale, census data is still processed from areal aggregates. Even if this was not a problem, there are physical barriers within the landscape that go against the view of population as continuous. Thus, several authors have cautioned against the use of surfaces when portraying census data.

> *"A surface is not an alternative for a choropleth map, because on the one hand it contains more information and on the other hand it contains less"(Dorling, 1995, p. 207).*

> *"It is extremely wise to avoid contour mapping of census data"(Rhind cited in Dorling, 1995, p. 207).*

Despite problems, surface representations are used increasingly for representing ethnic census data.

4.2.1.4. Proportional Point Symbols

A proportional point is a symbol that varies in proportion to the value that it represents. Although the symbol is usually a circle or some other standard geometric shape (e.g. triangle, square), it can be any point symbol. "There are two commonly accepted instances when graduated point mapping is selected by the cartographer: when data occur at points and when they are aggregated at points within areas. ... whenever the goal of the map is to show relative magnitudes of phenomena at specific locations, the graduated symbols form of mapping is appropriate" (Dent, 1993, p. 170).

Figure 4.20: Proportional Point Symbol Map of People of Single Italian Origin (1991), by Cenus Tract

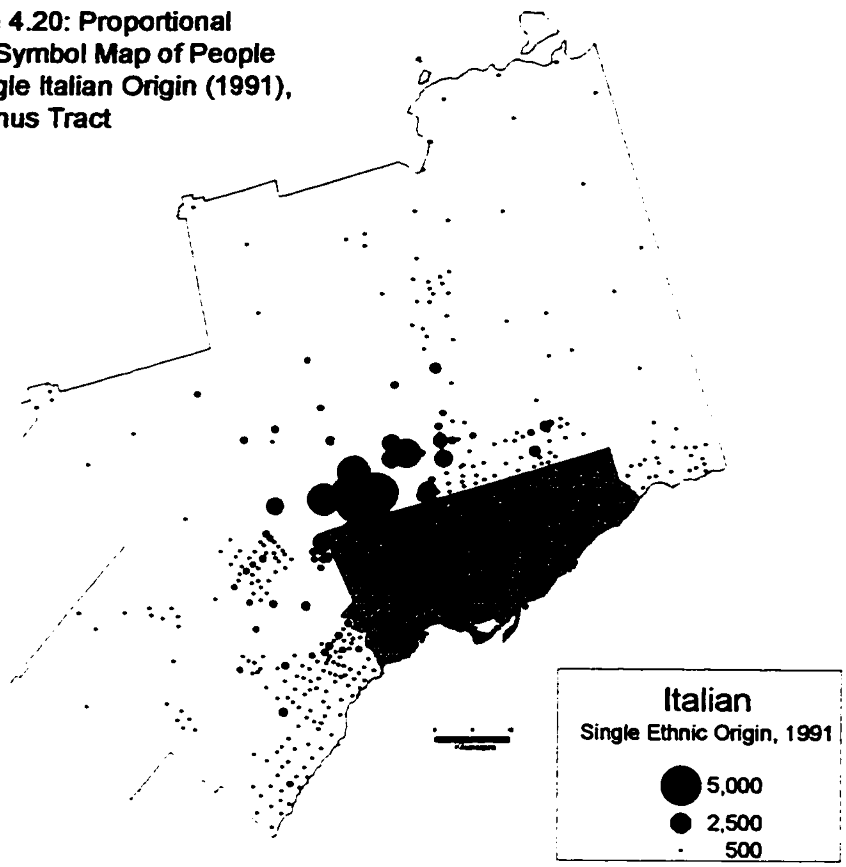
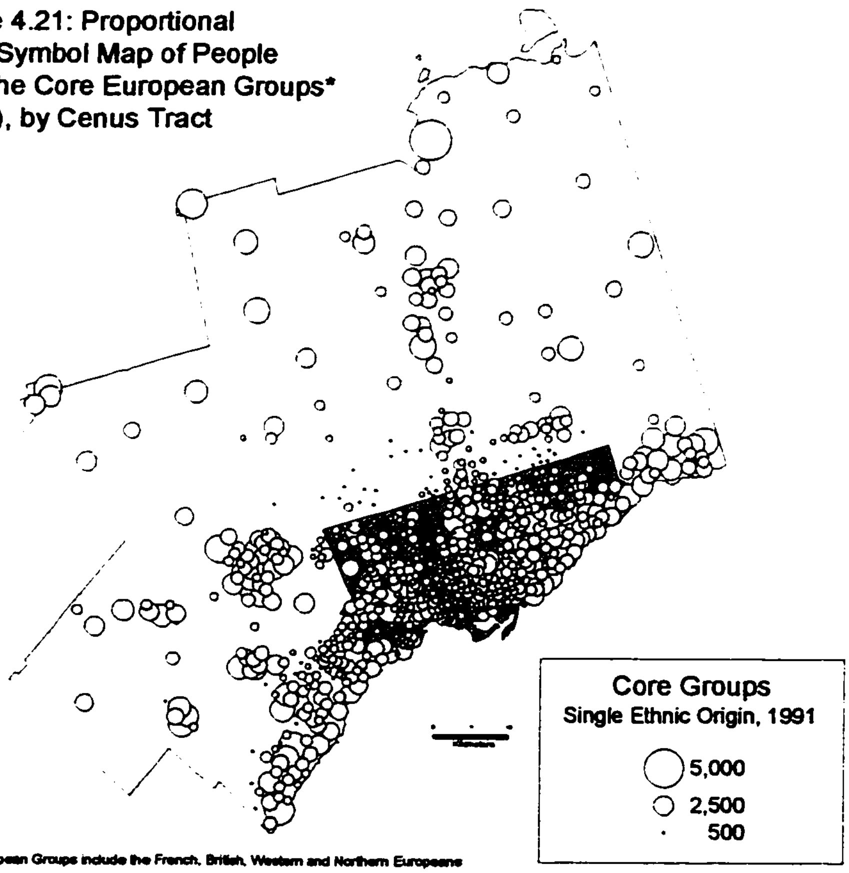


Figure 4.21: Proportional Point Symbol Map of People from the Core European Groups* (1991), by Cenus Tract



* The Core European Groups include the French, British, Western and Northern Europeans

With proportional symbols, one can choose between different symbols and different methods of scaling. The most common (for proportional circles) are absolute and psychological scaling. Psychological scaling does not scale by absolute value but by preconceived magnitude to account for the human tendency to underestimate differences in size. However, this option is not available in most standard mapping packages (e.g. MapInfo, ArcView, etc.).

Proportional point symbols have the advantage of being easy to produce and aesthetically pleasing. Also, proportional point symbols have the advantage of using census centroid files instead of boundary files. Although proportional symbols are very effective in showing phenomena that varies in intensity, they are not as effective in showing variables that remain uniform. For example, the absolute magnitude between census tracts is very minimal because census units, at each level of aggregation, are similar in population size. Census tracts ideally have populations between 2,500 and 8000, with a preferred average of 4,000. Therefore, proportional symbols are a poor choice for population density maps that use census units. However, proportional symbols are a good choice for showing the ethnic groups where the relative magnitude changes between census units.

Figure 4.20 and 4.21 show two proportional point symbol maps, one of the Italian group and one of the core European groups (French, British, Western and Northern Europeans). Figure 4.20 illustrates the varying distribution of the Italian group quite well. However, the proportional point symbol map of the core groups (Figure 4.21) is not as effective as the map of the Italian group because the population of the core groups is distributed more evenly across census tracts.

Another problem with proportional point symbol maps relates to the object class of proportional symbols. Ideally, proportional symbols represent data that is discrete and abrupt. This is generally not the case with census data. Despite disadvantages, however, proportional symbols remain a viable alternative for representing census data.

4.2.2. Multivariate Maps

Two multivariate maps have also been produced: a) proportional pie charts and b) nominal areal maps classed by dominant ethnic group.

4.2.2.1. Proportional Pie Charts

Proportional pie charts fall under the larger heading of proportional point symbols. Proportional symbols can be either univariate or multivariate. A common multivariate proportional symbol is the proportional pie chart. With the pie chart, the size of the circle is proportionate to the sum of the total population, and each slice represents a percentage of the whole.

Figure 4.22 provides an example of a proportional pie map. The census tract level of aggregation was used because it provides a neighbourhood level of detail. The enumeration level was not used because the number of points (centroids) created from enumeration areas would have been unreadable. Also, the scale had to be larger to accommodate extra information.

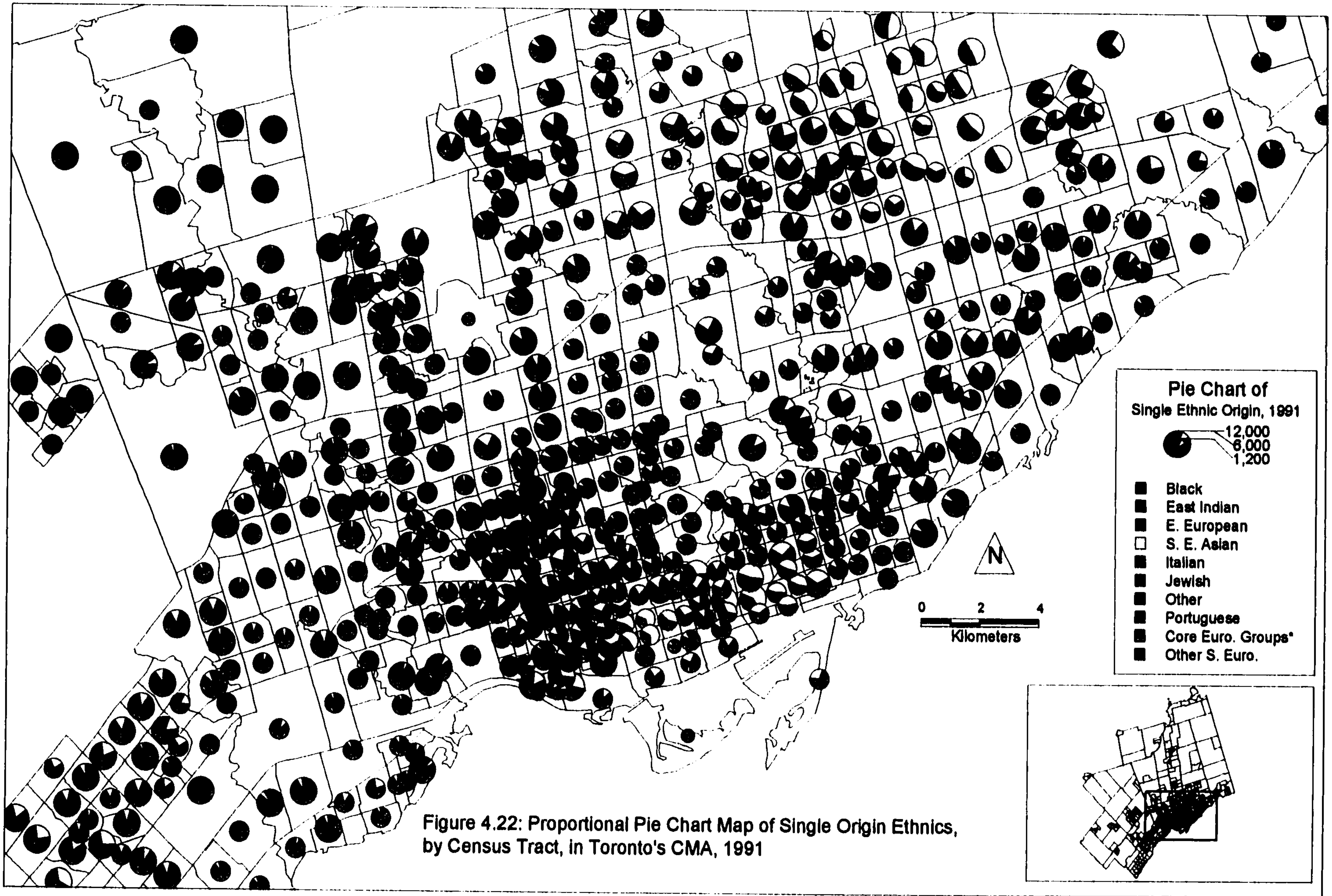


Figure 4.22: Proportional Pie Chart Map of Single Origin Ethnics, by Census Tract, in Toronto's CMA, 1991

* Core European Groups: British, French, Northern European, Western European

Proportional pie maps display the relative proportion of groups and the degree of homogeneity of spatial units. Because the proportional pie map displays the relative proportion of several groups, it graphically reveals the dominant ethnic groups. However, proportional pie maps are difficult to read when they hold too much information. "Bar charts, graphs and pyramids were originally designed to stand alone, and thus often contain enough complexity and detail as single entities, and too much when many are displayed simultaneously as map symbols" (Dorling, 1995, p. 197). "Symbol overload" (Dent, 1993), is when there are too many symbols or the symbols carry too much information and become unreadable or confusing to the reader. This is a common problem with proportional symbols, in general, and proportional pie charts, in particular. Notably, the pie map in Figure 4.22 would have been clearer if there had been fewer categories. However, for consistency, all groups within this study have been included. Also, a large number of groups is useful for showing how easily a map can become confusing (difficult to read). However, if fine detail is not required, the proportional pie map does a reasonable job of showing homogeneity, dominant ethnic groups and major ethnic enclaves.

Another problem with proportional pie charts is that the absolute magnitude between census tracts is very minimal. Therefore, most of the variation is in the slices of the pies and not in the size of the charts themselves. Also, the order of the variables in the chart affects the visual impression (Dorling, 1995), which is usually arbitrary.

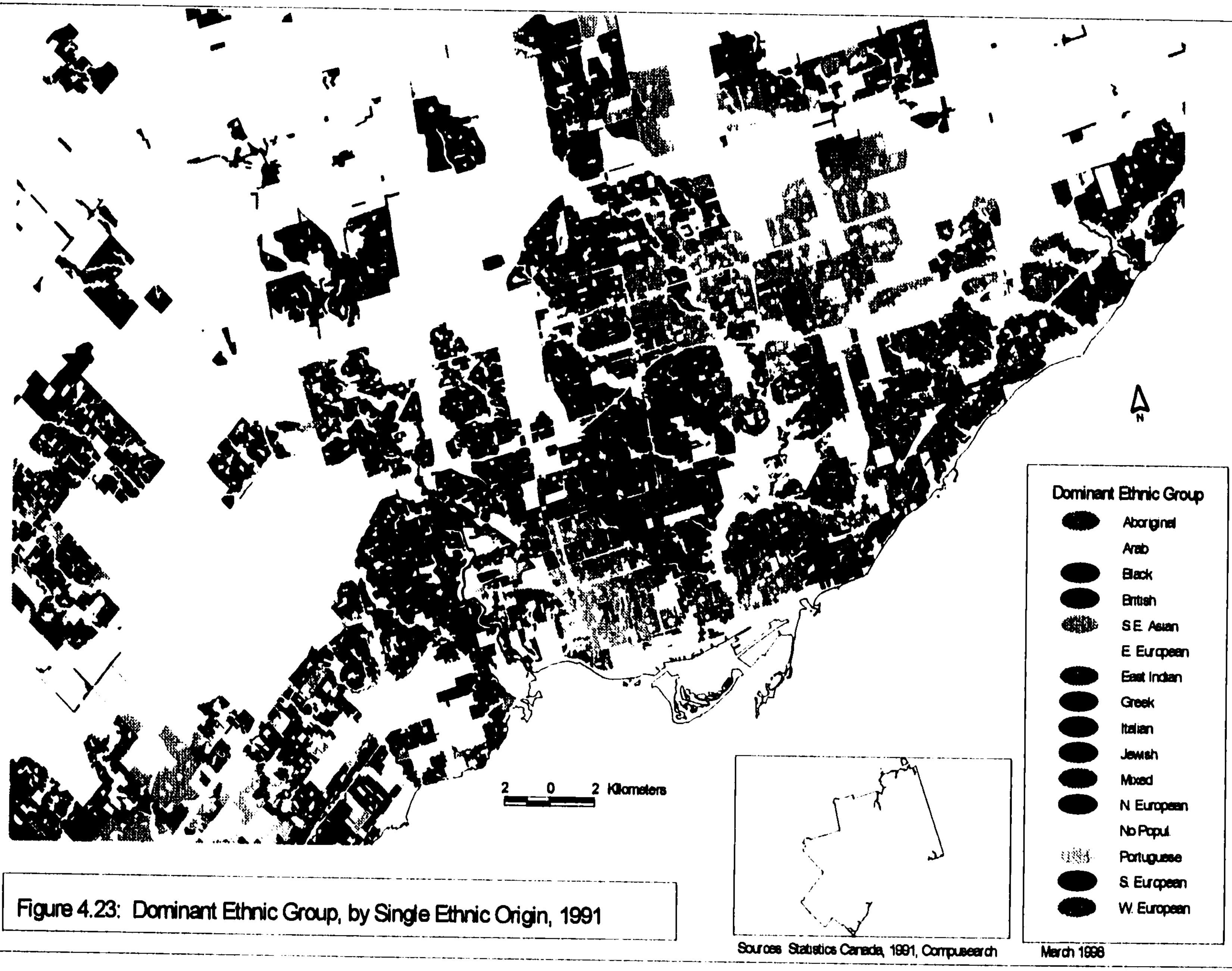


Figure 4.23: Dominant Ethnic Group, by Single Ethnic Origin, 1991

4.2.2.2. Nominal Areal Maps

Another way of representing ethnic data is to label the spatial units by their dominant ethnic group. This produces a nominal areal map. When a map displays nominal information, the units should be displayed in a way that does not imply a hierarchy or order. A hierarchy or order is usually achieved by changing the hue (or colour) instead of the saturation or value of the colour. Also, "some level of generalization is to be expected on maps with nominal distributions. ... Depending upon the scale of the map, however, these subareas may be omitted so that only the predominant ... type is indicated. If the number of subareas of different types is significant, the classification scheme should include suitable mixed categories" (Campbell, 1991, p. 212).

Figure 4.23 provides an example of a nominal areal map. Dasymetric enumeration units were used as the base spatial units. Similar to proportional pie charts, dominant areal maps allow the reader to see several groups at once. Also, they allow the reader to identify large ethnic enclaves. Furthermore, dominant areal maps are easier to read because the data has been processed and classified into dominant groups. This processing clarifies the cartographic message by reducing the amount of information within the map. Nominal areal maps, however, lose the homogeneity and proportional information that is portrayed in proportional pie maps. Also, a large number of categories can be confusing. Ideally there should have been less categories used in Figure 4.23; however, for consistency, all ethnic groups in this study have been included.

4.2.3. Cartograms and Statistical Maps

"The cartogram represents one level of abstraction beyond the land (choropleth) map - one level closer to the maps of social landscape we are aiming to see"
(Dorling, 1995, p. 175).

Cartograms and statistical maps have been put into the same category because both represent a higher level of abstraction.

4.2.3.1. Statistical Maps - Location Quotients

Location quotients are commonly used in economic geography. The location quotient compares the share of one activity with a benchmark activity. In this case, the proportion of group x is compared with a proportion of group y / or the total population. Location quotients are read in the following manner.

LQ > 1	more share of activity
LQ = 1	equal share of activity
LQ < 1	less share of activity

Location quotient values can be mapped to produce a statistical map. Figure 4.24 shows the relative proportion of the Portuguese group compared to the total population of the census unit. Therefore areas in the city that have a higher share of Portuguese are highlighted. Also, the major Portuguese enclaves are noticeable immediately. A Location Quotient map tells us more than a standard proportional map because it compares the distribution of one group with the distribution of another. Statistical maps offer the advantage of descriptive statistics as well as the ability to identify location and place.

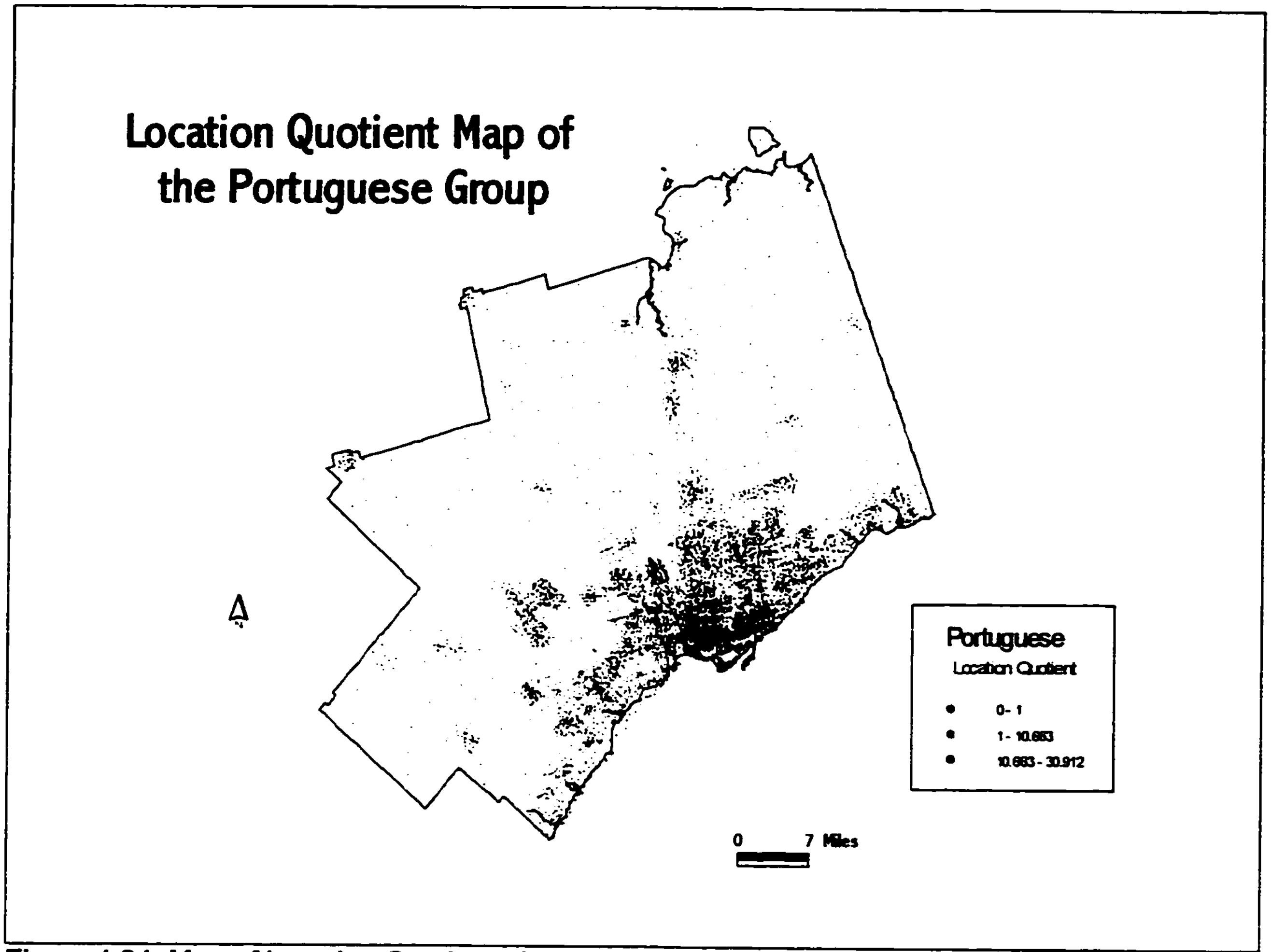


Figure 4.24: Map of Location Quotients for the Portuguese Group, by EA, Single Origin, 1991.

Figure 4.24 represents a simple version of a statistical map. Instead of using standard units, population weighted centroids were used. Using enumeration centroids gave each enumeration area the same visual impact. Notably, the indices produced in Chapter 3 could have also been mapped like the Location Quotient values.

4.2.3.2. Cartograms

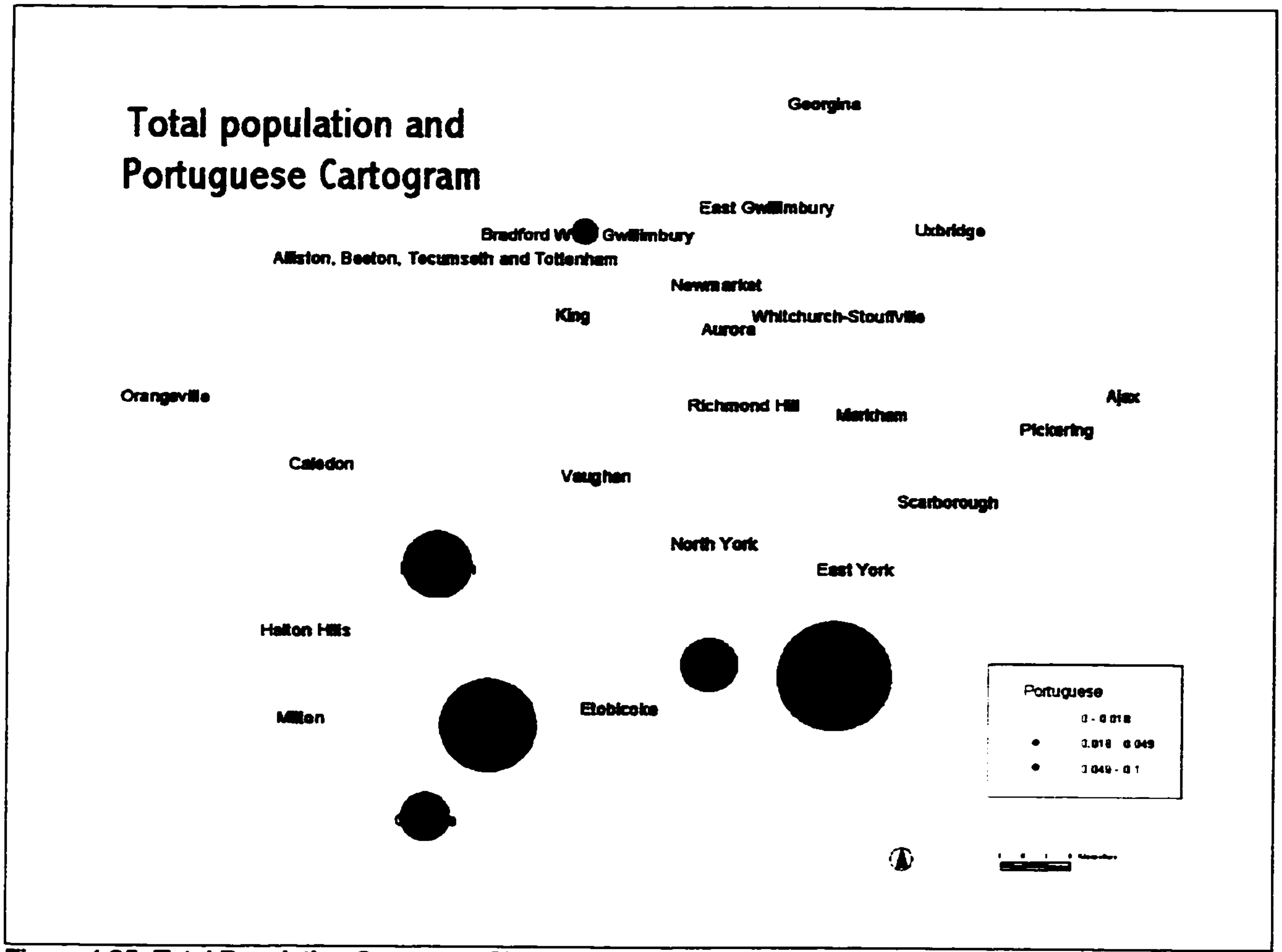


Figure 4.25: Total Population Cartogram Showing Proportion of Portuguese, Single Origin, 1991.

Cartograms are usually associated with value by area maps; however, cartograms can also represent maps with spatial transformations (Dent, 1993). Figure 4.25 shows a non-contiguous, population cartogram. With a population cartogram, "the arbitrary boundaries can be seen to have much less influence on the impression gained" (Dorling, 1995, p. 180), than if standard census boundaries are used. In Figure 4.25, the areas with the highest population (e.g.Toronto) are given the greatest visual emphasis. However, if this were a standard map of Toronto's CMA, the outer municipalities would have been given the greatest visual emphasis.

Because cartograms distort geographic space, they work best when the places depicted are well known by the reader. This is the main reason why cartograms are more suited to large, recognizable census units, e.g. CSDs (cities).

4.3. Summary

Although there are many ways of representing population data from the census, it is important to keep in mind the purpose of the map and the user. "Map making must always be viewed in the context of communication" (Dent, 1993, p. 2). Although this chapter has shown different representations, the question remains as to which representation is the best. Each representation, however, has certain advantages and disadvantages. Table 4.1 lists some of the main advantages and disadvantages of the different types of maps produced in this chapter.

Table 4.1: Types of Maps used for representing census population data

Choropeth	• easy to produce • shows data collection boundaries • identify areas that have a higher proportion of ethnics	• gives visual emphasis to large unpopulated areas	• method of classification can dramatically change the perceived level of segregation • should use ratio data
Dasymetric Choropleth	• uses populated land only • identify areas that have a higher proportion of ethnics	• time consuming • land use information can vary from user to user	• method of classification can dramatically change the perceived level of segregation • should use ratio data
Dominant Ethnic Group	• shows multiple groups	• time consuming	• large number of groups can be confusing to the reader • classification of areas open to misclassification
Dot Distribution	• able to visualize population density • of the same object class • does not place population in unpopulated areas	• time consuming	• size and value of the dots can dramatically change the perceived level of segregation
Proportional Pie Charts	• easy to produce • able to use centroid files instead of boundary files • shows multiple groups • can show homogeneity	• symbol overload	• large number of groups can be confusing to the reader
Proportional Symbols	• easy to produce • able to use centroid files instead of boundary files	• symbol overload	• only good for showing data that varies in intensity (eg. not good for showing total pop. in census tracts)
Surface	• asethically pleasing • able to visualize population density	• removed from the original data • can be very time consuming, depending upon type of interpolation	• many types of interpolation methods can produce many different results
Statistical	• can describe more complex spatial patterns and relationships	• can be time consuming, depending upon the Statistic	• moves closer to data analysis
Cartogram	• gives visual emphasis to places with large populations	• spatial transformations can be confusing to reader • can be time consuming, depending upon the type of cartogram	• has to be produced outside of standard mapping software

CHAPTER 5

CONCLUSION

5.1 Concluding Remarks

The advent of new technology has resulted in more numerous and detailed representations. We are now able to create representations that were previously not possible or too time consuming. However, these representations have not necessarily improved the way that we view segregation. We are still limited by the data and the amount of information that we can process. In addition, the truth and/or accuracy of current geographic representation comes into question with the authority and visual appeal afforded computer images. The allure and seductive quality of computer images placates and blinds the reader, e.g. he/she no longer questions the accuracy of representation. The reader also fails to question the truth of representation because of the legitimacy and authority afforded a mathematical, scientific image. However, the reader should always keep in mind that representation is an abstraction. As such, representation is partial, selective and subjective. Therefore, it can be said that all representation is but a partial rendering of the truth.

In Chapter 3, several indices were calculated at four standard levels of aggregation. Indices were calculated at different levels of aggregation to illustrate their sensitivity to the aggregation level and the spatial units used. Changing the level of aggregation was useful for highlighting the properties of the indices, revealing the spatial patterns of the ethnic groups, illustrating the sensitivity of the indices to different levels of scale and aggregation, and illustrating the affects of MAUP. However, different versions can be confusing to the reader. Notably, in this study, several observations were made. Several of the indices changed significantly when the level of aggregation was changed (see Chapter 3). However, since MAUP affects spatial patterns differently, the relative segregation standings were not always uniform. Therefore, the segregation rankings were often more revealing. Although there is no way of resolving the problems of arbitrary spatial boundaries and MAUP, there are some ways of making them less problematic. The following are ways of dealing with arbitrary spatial boundaries and MAUP when calculating indices.

- calculating indices at different levels of scale and aggregation;
- looking at the relative standing of indices;
- aggregating to more meaningful units.

Indices that used area in their calculations were also calculated using residential area only. The results were that both the RCO and ACE indices were reduced when the area was changed. This also illustrated the fluid nature of the indices. However, although the use of residential area was interesting, it was thought to be too time consuming for the average user.

In Chapter 4, several types of maps were produced. These maps were produced to illustrate some of the alternatives available for representing ethnic census data. However, there are numerous problems associated with census data. Although there is no way of alleviating all of these problems, there are some ways of reducing their visual impact. The following maps are more suited for dealing with arbitrary spatial boundaries:

- dasymetric maps;
- dot distribution maps;
- cartograms;
- maps that use centroid information.

With maps, the ways of dealing with MAUP include:

- producing different versions of the same map by changing the spatial units and level of aggregation to illustrate the effects of MAUP;
- using the smallest spatial units, e.g. dot distribution maps;
- using appropriate mapping techniques.

Maps are limited by both the data and a visual component. The visual is restrictive because the eye can only process a limited amount of information. Generally, the more places there are, the less information that can be displayed about each place. Therefore, different types of maps are more suited to different levels of aggregation. For example, at a metropolitan or city level of scale, dot distribution and surface maps are more suited to the EA level of aggregation than the choropleth or proportional point symbol maps. This is because it is not as important to see the individual spatial units as to see the general patterns that emerge with dot distribution and surface maps (e.g. dots coalesce). The choropleth and proportional point symbol maps are more suited to the census tract level of aggregation. At the enumeration level, many of the small spatial units would not be visible on a choropleth map, and too many proportional symbols would be produced and overlap on a proportional point symbol map. Proportional pie charts and bar charts are more suited to the larger spatial units (e.g. FED or CSD) because they hold more information. Cartograms are generally more suited to large, recognizable spatial units (e.g. CSD's – cities) because of their spatial transformation. Notably, these are general classifications to which many exceptions exist.

5.1.1. Suggestions for Future Research

This study has touched upon some of the possibilities for using census data to represent ethnic residential segregation. However, possibilities for future research include:

Indices:

- calculating indices on other cities within Canada;
    - calculating indices at different levels of aggregation;
    - calculating indices for different periods of time;
- calculating indices at the same level of aggregation but using different boundaries —> to aggregate up to meaningful units;
- calculating ACE index using criteria other than the CBD (e.g. closeness to subway or public transportation).

Maps:

- producing maps that use the index values calculated in Chapter 3;
- expert systems for mapping;
- seeing if the perception of ethnic enclaves (mental maps) coincides with actual concentrations of ethnic and racial groups.

Further research could also include a look at the audience and the purpose of maps and indices.

Appendix

```
10 PRINT "*** Lorenz Curves ***"
15 PRINT "--------------------"
20 DIM X(6000), R(6000)
30 PRINT "HOW MANY AREAS ";
40 INPUT N
50 LET X3 = 0!
70 FOR I = 1 TO N
75 PRINT "CASE "; I; " = ";
80 INPUT X(I)
95 LET R(I) = I
100 LET X3 = X3 + X(I)
120 NEXT I
130 FOR I = 1 TO N
200 LET X1 = X(I)
225 LET R1 = R(I)
300 NEXT I
310 PRINT "OUTPUT TABLE WITH CUMULANTS X2"
320 PRINT "CASE", "    X"
330 LET X2 = 0!
350 FOR I = 1 TO N
360 LET X2 = X2 + X(I) / X3
380 PRINT R(I), X2
390 NEXT I
400 SYSTEM
```

References

Anderson, Richard H. **Theory and Research: Approaches in Sociology.** Department of Sociology, University of Colorado at Denver, Colorado, 1996.

Armhein, C., Reynolds, H., Using the Getis Statistic to Explore Aggregation Effects in Metropolitan Toronto Census Data. in **The Canadian Geographer 41**, no 2, Canadian Associations of Geographer, 1997, pp 137-149.

Aron, A.; Aron, E. N., **Statistics for Psychology.** Prentice Hall, Englewood Cliffs, 1994.

Barber, G., **Elementary Statistics for Geographers.** The Gilford Press, New York, 1988.

Bertin, J., **Graphics and graphic information processing.** Walter de Gruyter, Berlin, 1981.

Blakemore, Michael, in Gregory, D.; Smith, D. (eds), **The Dictionary of Human Geography, Third Edition**. Blackwell Publishers Ltd., Oxford, 1994.

Blau, Peter M., **Inequality and Hetergeneity: A Primitive Theory of Social Structure.** Free Press, 1977.

Burrough, P.A., **Principles of Geographic Information Systems for Land Resources Assessment**. Oxford University Press, Oxford, 1986.

Bracken, I., Martin, D., The generation of spatial population distributions form census centroid data, in **Environment and Planning A.** 21, 1988, pp 537-543.

Bracken, I., Martin, D., Linkage of the 1981 and 1991 UK Census using Surface Modelling Concepts, in **Environment and Planning A.** 27, 1995, p 382.

Breton, R., Isaljiw, W., Kalbach, W., Reitz, J., **Ethnic Identity and Equality: Varieties of Experience in a Canadian City**. University of Toronto Press, Toronto, 1990.

Bourne, L., **Social Inequalities, Polarization and the redistribution of Income within Cities: A Canadian Example.** AAG Paper, Charlotte, NC, 1996.

Brunsdon, C., Anaysis of univerate census data. in Openshaw, Stan., ed. **The Census Users' Handbook.** John Wiley and Sons, New York, 1995.
Campbell, J., **Map Use and Analysis, second edition.** Wm. C. Brown Publishers, Dubuque, 1991.

Charlton, M. Rao, O., Carver, S., GIS and the Census. in Openshaw, Stan., ed. **The Census Users' Handbook.** John Wiley and Sons, New York, 1995, pp 133-165 .

Chrisman, N. R., Modeling error in overlaid categorical maps, in Goodchild, M. F., Gopal, S., eds., **Accuracy of Spatial Databases**. Taylor & Francis Ltd., London, 1989, pp 20-34.

Cloke, Paul J., The Differences of Postmodern Human Geography, **Approaching Human Geography, An Introduction to Contemporary Theoretical Debates**. Guilford Press, New York, 1991, pp 170-205.

Cortese, C. R., Falk, R. F., Cohen, J. C., Further Consideration on the Methodological Analysis of Segregation Indices. in **American Sociological Review.** 41, 1976, pp 630-637.

Cosgrove, Denis, Environmental Thought and Action: Pre-Modern and Post-Modern, in **Trans. Inst. Br. Geog.** NS, 15: ISSN:0020-2754, pp 344-58.

Dent, B. D., **Cartography Thematic Map Design**. Third Edition. WM. C, Brown Publishers, Dubuque, 1993.

Denton, Nancy, Are African American Still Hypersegregated? in Bullard, R.D.; Grigsby, J. E.; Lee, C, **Residential Apartheid: The American Legacy.** CAAS Publishing, Los Angeles, 1994, pp 49-81.

Dorling, D., Visualizing the 1991 Census. in Openshaw, Stan, ed. **The Census Users' Handbook.** John Wiley and Sons, New York, 1995, pp 167-211.

Duncan, O. D., Duncan, B., A Methodological Analysis of Segregation Indices. in **Population Studies.** 11, 1955, pp 210-217.

Duncan, J., Ley D., **Place/Culture/Representation.** Routledge, New York, 1993.

Eastman, J.R., Kyem, P.K., Tolerano, J., Weigen, J., **GIS and Decision Making**, UNITAR European Office, Geneva, 1993.

Fellmann, J., Getis, A., Getis, J., **Human Geography: Landscapes of Human Activities, Fourth Edition.** Wm. C. Brown Publishers, Dubuque, 1990.

Godlewska, Anne, Map, Text and Image. The Mentality of Enlightenment Conquerors: A New Look at the Description de l"Egypte in **Trans. Inst. Br. Georg.**, No. 20, 1995, ISSN-0020-2754, pp 5-28.

Goodchild, M. F., Anselin, L., Deichmann, U., A framework for the areal interpolaiton of socioeconomic data in **Environment and Planning A**, 25, 1993, pp 383-397.

Gregory, D., Smith, D., Johnson, R.J. (eds), **The Dictionary of Human Geography, Second Edition.** Blackwell Publishers Ltd., Oxford, 1986.

Gregory, D.; Smith, D. (eds), **The Dictionary of Human Geography, Third Edition**. Blackwell Publishers Ltd., Oxford, 1994.

Grigsby, J. E., African American Mobility and Residential Quality in Los Angeles, in Bullard, R.D.; Grigsby, J. E.; Lee, C, ; **Residential Apartheid: The American Legacy.** CAAS Publishing, Los Angeles, 1994.

Halli, Shiva, S., Frank Trovato and Leo Driedger, eds., **Ethnic Demography: Canadian Immigrant, Racial and Cultural Variations**. Carleton University Press, Ottawa, Ont., 1990.

Harley, J.B., Deconstructing the Map, in **Cartographica**, Vol. 26, No. 2, Summer, 1989, pp 1-20.

Harvey, D., **The Condition of Postmodernity, An Enquiry Into the Origins of Cultural Change.** Oxford: Blackwell, 1989.

Hecht, A.; Sharpe, R.; Wong, A. **Ethnicity and Well-Being in Central Canada: The Case of Ontario and Toronto (Kanada-Projekt II).** Im Selbstverlag des Geographischen der Universitat Marburg, Marburg, 1983.

Hiebert, D., Canadian Immigration: Policy, Politics, Geography, in **The Canadian Geographer 38.** no 3, Canadian Associations of Geographer, 1994, pp 254-258.

Hunt, L. **Theoretical and Empirical Examinations of Spatial Scale and Aggregation Effects on the Principal Axis Factoring Technique when the Observations are Areal Units.** (unpublished M.A. Thesis), Wilfrid Laurier University, Waterloo, 1993.

Jackson, P., **Ethnicity and Boundaries: 'Ordered Segmentation' in Urban Neighbourhoods.** Oxford Press, Oxford, 1980.

Jackson, Peter; Smith, Susan J., **Exploring Social Geography.** George Allen & Unwin Ltd., London, 1984.

James, David R; & Tauber, Karl E., Measures of Segregation, in **Sociological Methodology 1985.** Jossey-Bass, 1985, pp 1-32.

Langford, M. & Unwin, D. J., Generating and mapping population density surfaces within a geographical system. in **The Cartographic Journal,** 31, 1994, 21-26.

Laurini, R., Thompson, D., **Fundamentals of Spatial Information Systems.** Academic Press, San Diego, California, 1992.

Ley, David, Introduction: Representing the Place of Culture. In Duncan, J., & Ley, D. (eds.), **Place/Culture/Representation.** Routledge, London, 1993, p 4.

Lieberson, Stanley; Carter, Donna R., Temportal Changes and Urban Differences in Residential Segregation: A Reconsideration, in **American Journal of Sociology**, Volume 88 number 2, Chicago, 1988.

MacEachren, A.M., **Some Truth With Maps: A Primer on Symbolization and Design**. AAG, Washington, 1994.

Mark, D. M., Csillag, F., The Nature of boundaries in 'area-class' maps, in **Cartographica**. 26(1), 1989, pp 65-78.

Martin, D., **Geographic Information Systems and Their Social Economic Applications.** Routledge, London, 1991a.

Martin, D., Mapping population data from zone centroid location, in **Transactions of the Institute of British Geographers**. New Series, 14(1), 1991b, pp 90-97.

Martin, D., **Geographic Information Systems Social Economic Applications, Second Edition.** Routledge, London, 1996.

Massey, D., Ethnic Residential Segregation: A Theoretical Synthesis and Empirical Review, in **Sociology and Social Research 69.** 1985, pp 315-50.

Massey, Douglas S., Denton, Nancy A., The Dimensions of Residential Segregation, in **Social Forces**. Volume 67:2, December 1988.

McCleary, C. F., An effective graphic 'vocabulary', in **IEEE Computer Graphics and Applications**, March, 1983, pp 46-53.

Morrison, J. L., Applied cartographic communitcation: Map symbolization for atlases. in Board, C., ed., **New insights in cartographic communication, Cartographica Monograph 31**, 1984, pp 44-84.

Murray, Jeffrey S., Maps that Deceive. in **Canadian Geographic**, May/June 1992.
Openshaw, Stan, The Modifiable Areal Unit Problem, in **Catmog 38**. Geo-Abstracts, Norwich, 1984a.

Openshaw, Stan, Ecological fallacies and the analysis of areal census data, in **Environment and Planning A** 16, 1984b, pp 17-31.

Openshaw, Stan., Learning to live with error in spatial databases, in Goodchild, M. F., Gopal, S., eds., **Accuracy of Spatial Databases**. Taylor & Francis Ltd., London, 1989, pp 263-276.

Openshaw, Stan., (ed) **The Census Users' Handbook**. John Wiley and Sons, New York, 1995.

Perry, Greg, **Qbasic By Example**. Que Corporation, Indianapolis, 1993.

Pickles, John, Representations in an Electronic Age, in Pickles, John, ed., **Ground Truth: The Social Implications of Geographic Information Systems**. The Guilford Press, New York, 1995, p 12.

Ray, B. Immigrant Settlement and Housing in Metropolitan Toronto, in **The Canadian Geographer 38**, no 3, Canadian Association of Geographers, 1994, pp 262-265.

Smith, Susan J., **The Politics of 'Race' and Residence: Citizenship, Segregation and White Supremacy in Britain.** Polity Press, Oxford, 1989.

Stearns, L. B., Logan, J. R., Measuring Segregation: Three Dimensions, Three Measures. in **Urban Affairs Quarterly** 22, 1986, pp 124-150.

Tomlin, C. D., **Geographic Information Systems and Cartographic Modelling**. Prentice Hall, Engelwood Cliffs, NJ., 1991.

Tufte, E., **Envisioning Information**. Graphic Press, Cheshire, 1990.

Turnbull, David, **Maps and Territories: Science is An Atlas**. The University of Chicago Press, Chicago, 1993 ed.

Unwin, D. J., Dawson, J. A., **Computer Programming for Geographers.** Longman Inc., New York, 1985.

Veregin, H., Error modeling for the map overlay operation, in Goodchild, M. F., Gopal, S., eds., **Accuracy of Spatial Databases**. Taylor & Francis Ltd., London, 1989, pp 4-18.

Weiher, Gregory R., **The Fractured Metropolis: Political Fragmentation and Metropolitan Segregation**. State University of New York Press, Albany, 1991.

White, M. J., The Measurement of Spatial Segregation. in **American Journal of Sociology**, 88, 1983, pp 1008-1019.

Wong, David W.S. Spatial Dependency of Segregation Indices, in **The Canadian Geographer 41**, no 2, Canadian Associations of Geographer, 1997, pp 128-36.

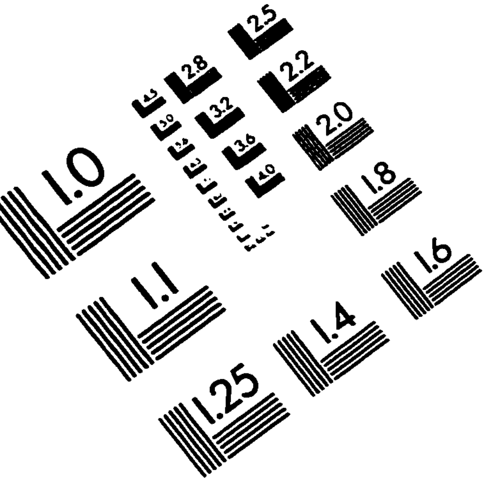
Yeates, M. **The North American City.** Harper Collins, New York, 1990.

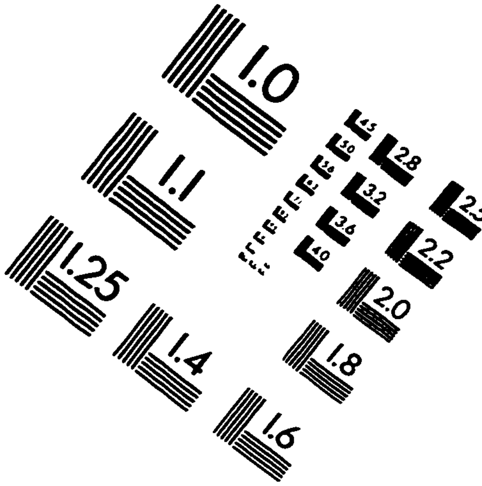
. . . . . **1991 Census Dictionary.** Statistics Canada, Ottawa,1992.

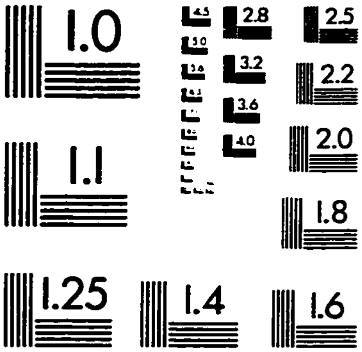
. . . . .Challenges of Measuring an Ethnic World: Science Politics and Reality, in **Proceedings of the Joint Canada - U.S. Conference on the Measurement of Ethnicity April 1-3,1992.** Statistics Canada, Ottawa | U.S. Department of Commerce.

. . . . . **Research Infopac Series,** City of Toronto Planning and Development Department, Toronto, 1994.

IMAGE EVALUATION TEST TARGET (QA-3)

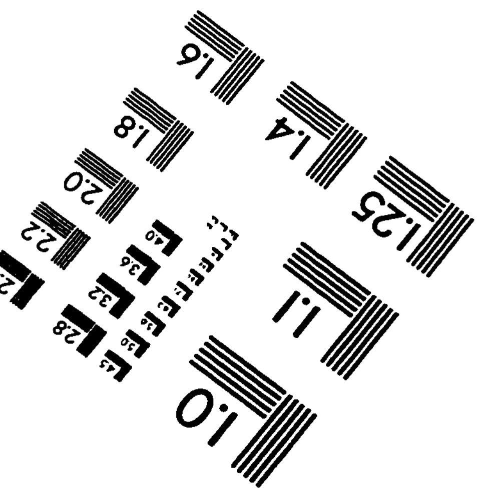


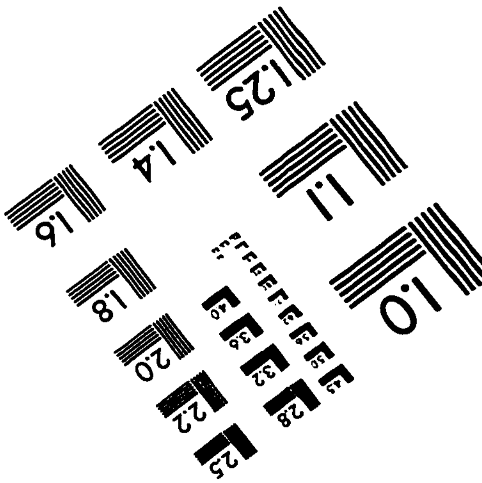




150mm

6″





APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989