# A NEW CENSUS GEOGRAPHY FOR TURKEY USING GEOGRAPHIC INFORMATION SYSTEMS A CASE STUDY ON ÇANKAYA DISTRICT, ANKARA

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

CEM KIRLANGIÇOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE
IN
GEODETIC AND GEOGRAPHIC INFORMATION TECHNOLOGIES

JUNE 2005

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan Özgen
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Zuhal Akyürek
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc.Prof.Dr. Oğuz Işık
Supervisor

**Examining Committee Members**

| | | |
|---|---|---|
| Assist.Prof.Dr. Zuhal AKYÜREK | (METU, GGIT) | |
| Assoc.Prof.Dr. Oğuz IŞIK | (METU, CRP) | |
| Assoc.Prof.Dr. Nurünnisa USUL | (METU, CE) | |
| Assoc.Prof.Dr. M. Melih PINARCIOĞLU | (METU, CRP) | |
| Assoc.Prof.Dr. Murat GÜVENÇ | (METU, CRP) | |

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name, Last name : Cem Kırlangıçoğlu

Signature :

# ABSTRACT


A NEW CENSUS GEOGRAPHY FOR TURKEY
USING GEOGRAPHIC INFORMATION SYSTEMS:
A CASE STUDY ON ÇANKAYA, ANKARA

Kırlangıçoğlu, Cem
M.S., Department of Geodetic & Geographic Information Technologies
Supervisor: Assoc.Prof.Dr. Oğuz Işık

June 2005, 123 pages

Today, population census results are not only fundamental for population counting but also for providing various information to many organizations and people who make research and take decisions about human-related issues. However, statistics produced on the basis of administrative divisions in Turkey are inadequate to meet the needs of most data users.

In this study, a new census geography is aimed for Turkey using Geographic Information Systems and Multi Criteria Decision Making methodologies. This new census geography is for statistical purposes only, independent from legal boundaries and generated by the concept of small area statistics. New rules and methodologies are created by taking the United States and the United Kingdom systems as models to reach to the main aim, and then they are applied on a case study area, Çankaya District in Ankara. Through this application process, firstly all the collected graphical and non-graphical raw data are geo-referenced and combined in a common geodatabase. Secondly, this geodatabase is used to understand the differentiation of quality of life indicators across the case study area. This pattern is then used to draw the boundaries of small statistical units of the new census geography in terms of the previously defined population sizes. In conclusion, a nationwide standard census geography hierarchy, which ranges between national level at the top and block level at the bottom, is proposed for use in 2010 Population Census and afterwards.

Keywords: Census Geography, Geographic Information Systems (GIS), Small Area Statistics, Multi Criteria Decision Making (MCDM), Turkey.

# ÖZ


COĞRAFİ BİLGİ SİSTEMLERİ DESTEĞİYLE
TÜRKİYE İÇİN YENİ BİR SAYIM COĞRAFYASI,
ÇALIŞMA ALANI: ÇANKAYA İLÇESİ, ANKARA

Kırlangıçoğlu, Cem
M.S., Jeodezi ve Coğrafi Bilgi Teknolojileri Ana Bilim Dalı
Danışman: Doç. Dr. Oğuz Işık

Haziran 2005, 123 sayfa

Günümüzde nüfus sayımları sadece nüfusu saymaktan öteye geçmiş, insana ve topluma dair bilgiye ihtiyaç duyan pek çok farklı kurum ve araştırmacının temel bilgi kaynağı haline gelmişlerdir. Fakat Türkiye'de idari sınırlar baz alınarak üretilen sayım tabanlı istatistikler çoğu istatistiki veri kullanıcısının ihtiyaçlarına yeterince cevap veremez durumdadır.

Bu çalışmanın amacı, Coğrafi Bilgi Sistemleri'nin ve Çok Kriterli Karar Verme yöntemlerinin de yardımıyla, Türkiye için yeni bir sayım coğrafyası önermektir. Bu yeni sayım coğrafyası sadece istatistik amaçlı, idari sınırlardan bağımsız ve küçük alan istatistiği anafikrine dayalı olarak tasarlanmıştır. Gerekli kural ve metodolojiler, Amerika Birleşik Devletleri ve Birleşik Krallık modellerinden yola çıkılarak tasarlanmışlar ve Ankara'nın Çankaya ilçesinde yapılan örnek çalışmada uygulanmışlardır. Bu uygulama sürecinde; ilk olarak toplanan grafik ve grafik olmayan ham veriler birbirleriyle ilişkilendirilmiş ve ortak bir veritabanında bir araya getirilmişler, ardından da bu veritabanı yardımıyla çalışma alanındaki yaşam kalitesi farklılaşmalarının deseni anlaşılmaya çalışılmıştır. Bir sonraki adım olarak ta, bu desen dahilindeki mümkün olduğunca homojen alanlar, belirlenmiş nüfus kriterleine göre bölgelenmiş ve yeni sayım coğrafyasının istatistiki birimlerinin yaratılmasında kullanılmışlardır. Sonuç olarak, en üst seviyede Türkiye’nin ve en alt seviyede de yapı adasının bulunduğu, 2010 genel nüfus sayımı ve sonrasında kullanılabilecek, ulusal kapsamlı bir standart sayım coğrafyası hiyerarşisi önerilmiştir.

Anahtar Kelimeler: Sayım Coğrafyası, Coğrafi Bilgi Sistemleri (CBS), Küçük Alan İstatistiği, Çok Kriterli Karar Verme, Türkiye.

To My Family…

To My Love …

## ACKNOWLEDGMENTS

I gratefully thank to Assoc. Prof. Dr. Oğuz IŞIK, my supervisor, for his valuable comments on my thesis and his great influence to my work. He has reminded me to stop running and first to think.

I would also like to thank all my lecturers; Assist. Prof. Dr. Zuhal Akyürek, Assoc. Prof. Dr. Nurünnisa Usul, Assist. Prof. Dr. Şebnem Düzgün, Prof. Dr. Vedat Toprak, Assist. Prof. Dr. Mustafa Türker, Prof. Dr. Adnan Yazıcı, Assist. Prof. Dr. M. Onur Karslıoğlu, Onur Paşaoğlu, Levent Ucuzal, and Tuncay Küçükpehlivan.

I would like to thank my best friends at METU; A. Kürşat Kurtar, Eda Erkin, Y. Kazım Ölçer, and Y. Kerem İnceçayır. Those long hours in GGIT lab will not be forgotten. Also thanks to my other friends; Çağatay Yamak, Emrah Söylemez, Gözde Kaya, Özlem Şevik, Volkan Kepoğlu, Ece Aksoy, Abdülvahit Torun, Kıvanç Ertuğay, and all GGIT assistants for their warm friendships.

I would also like to acknowledge Kıvanç Ertuğay for the basic graphical data support, and to the staff of the State Institute of Statistics (SIS) of Turkey for their non-graphical data support about the case study area.

I am greatly indebted to my mother Sevim Kırlangıçoğlu, my father Yener Kırlangıçoğlu, and my two elder brothers Ahmet and Nuri Kırlangıçoğlu because of their never-ending support.

Final thanks go to the God, who has granted me such a magnificent fiancée, Duygu Karadeniz.

**TABLE OF CONTENTS**

# LIST OF TABLES

TABLES

# LIST OF FIGURES

FIGURES

## LIST OF ABBREVIATIONS

AHP: Analytic Hierarchy Process

BG: Block Group

CT: Census Tract

DBMS: Database Management System

EU: European Union

GIS: Geographic Information Systems

GUI: Graphical User Interface

LAU: Local Administrative Units

MADM: Multi Attribute Decision Making

MCDM: Multi Criteria Decision Making

MERNIS: Central Population Management System

METU: Middle East Technical University

NUTS: Nomenclature Units for Territorial Statistics

RS: Remote Sensing

SAS: Small Area Statistics

SAW: Simple Additive Weighting

SIS: State Institute of Statistics

SPO: State Planning Organization

SSA: System of Small Areas

TAKBIS: Turkish Land Registry and Cadastre Information System

TBMM: The Turkish Grand National Assembly

UK: United Kingdom

UN: United Nations

US: United States

WWW: World Wide Web

# CHAPTER 1

# INTRODUCTION

The Turkish Republic will conduct its 15th population census in the year of 2010. The purpose of this population census will be "…to determine completely and correctly the size, the distribution by the *administrative* division, and the demographic, social and economic characteristics of the population within the boundaries of the country on the application date of the census." (SIS, 2005)

Today, census results are not only fundamental for population counting but also for providing information to many organizations and people who make researches and decisions about many human-related issues. "These include matters of public relevance such as health and education, transportation planning and community services, and private concerns - such as siting of businesses, housing, consumer marketing, and economic strategies*" (*Edmonston, 1999).

"Recent increasing importance of information and knowledge systems in the new process oriented methods adopted in the production, distribution and consumption of goods and services in the societies caused the demand for a new geographical base for statistics" (Backer et al., 2002). From this point of view; is the geographic base of censuses (administrative division) adequate to answer the needs of Turkey? The answer is 'No' because of several reasons.

First of all, administrative boundaries are always subject to change because of political or population trends. As a matter of fact, administrative divisions of Turkey have been changed so many times; and these alterations cause serious problems for statistical data users by making time-series comparisons almost impossible. Researchers may not have accurate results from their studies; formulation of regional policies and development plans may cause unsatisfactory results; and goods and services may not be conveyed to the needy areas. Therefore, use of relatively *unchangeable statistical units* is needed to improve the current statistical system of Turkey.

The second drawback of legal boundaries is their non-standardized structure. It is obvious that almost none of the administrative units are comparable to each other in terms of population numbers. For example, the population of a quarter may be 30-40 times larger than another in the same district. This condition prevents making good judgments among administrative units, which are also statistical areas, and makes them incomparable to each other. Therefore, Turkish statistical system needs standardized and comparable statistical units.

The third considerable disadvantage is the possibly misleading effects of the population censuses. Yetik (2003) mentions that "the majority of statistics, being developed, are on the basis of the cities, which are the biggest geographical unit of administrative classifications and very few of them are based on districts." The provinces and districts have relatively big sizes in respect to population; and when statistics are produced for these divisions, the specific characteristics of local inhabitants are ignored. As a result, data users see only the average, not the real. Therefore, Turkish statistical system needs homogeneous statistical units to the possible extent to reduce such kinds of errors.

The number of drawbacks may be increased, but these three reasons are adequate to show the disadvantages of the geographic bases of the current system. It is clear that, administrative areas have been designed for administrative purposes only, not for statistical purposes. Thus, starting out from the drawbacks of the current system, **the main aim of this thesis** is to create a new and only statistical purposed '*Census Geography'* for Turkey. Census Geography is "…a collective term referring to the geographic entities used in data collection and tabulation operations, including their relationships to one another." (US Census Bureau, 2002, p.9) This new system will not completely substitute the current statistical constitution and administrative division, rather it must be thought as a complementary to the present system.

There are many steps throughout the thesis to reach to the main aim. First step includes a literature review mainly about the concept of ‘census geography’. This literature review also showed that there is almost no study concerning with the census geography needs of Turkey. There are some studies that take the first step about those subjects and demonstrate some kind of GIS usage in such kind of studies, but none of them propose concrete answers to the question of how a census geography should be in Turkey.

The literature survey also helped to understand the present census-based statistical system, administrative system and already in use census geography of Turkey. Of course, it would be very hard to interpret the current system without knowing the others in the world. Therefore international applications and census geographies were examined. Then, the countries with the most developed census geography, the USA and the UK, were taken as the main models because of their reliable, respected, settled and well-working statistical systems. At this point, it was impossible to take their system exactly the same and to apply it in Turkey because of structural differentiations, which will be explained in Chapter 2 and Chapter 3. So, the rules and methodologies of the new census geography have been proposed for Turkey after the necessary adaptation operations on the US and the UK systems.

The main point of those rules is to establish and identify the new census geography based on the concept of *small area statistics*; which requires basically seven fundamental requirements for the census geography units; homogeneity, functional integrity, compactness, continuity and consistency, population equity with same level of statistical areas, easy identification, and historical comparability. These concepts will be explained in detail in Chapter 3.

In the next step, in an attempt to decide what data to use, the physical and conceptual data models have been designed in the light of the created rules and methodologies. Then, the case study area (Çankaya District) was selected and data collection period began. Here, the major data source was the Prime Ministry State Institute of Statistics (SIS), whose task is to collect, evaluate and publish every kind of statistical data about economic, social, cultural and demographic structure of the nation. Of course, all the gathered data, some of which were in different formats, different coordinate systems, different characteristics etc. needed some manipulations like data conversion, editing, integration and also data generation. The purpose of these manipulations was to combine all the data in a common database to create a common language among them.

After collecting and combining all the data, the next step includes effective use of them in a Multi Attribute Decision Making procedure, Analytic Hierarchy Process (AHP), to estimate the distribution of 'quality of life' degrees in the area. Here the aim is to determine the boundaries of homogeneous areas to the possible extent, and include the inhabitants with similar types of living conditions into the same geographic units. For that purpose, basically 10 human development indicators were used in the AHP procedure by the help of a computer program running on Visual Basic platform. Then, a sensitivity analysis was performed to examine the sensitivity of the results to changes in the priorities of the indicators.

After having satisfactory results in the sensitivity analysis, the last step was to draw the final boundaries of the statistical units in a hierarchical organization, which is necessary to make the new system serving different kinds of data needs of various user groups such as governmental organizations, universities, private sectors etc. Here the boundaries were drawn with the help of the homogeneous areas found by the computer program, seven fundamental requirements defined above, and housing characteristics of the urban parts. Through this process, the 'Districting' extension of ArcGIS was used to define the boundaries of census geography units in terms of defined population sizes.

Finally unique region codes were given to the defined census geography units to make their names distinguishable from each other in practical applications.

Geographic Information System (GIS) was used as the main tool through the whole process of the study because of its strong abilities in complex data operations. A GIS is briefly “an organized collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information” (WebGIS, 2005). GIS, when properly used, allows end users to produce better decisions faster, and provide its users with the ability to make sense out of data that may otherwise be overlooked.

After this Introduction chapter, the thesis continues with Chapter 2, which makes descriptions of GIS and its uses in population censuses. Then it states the importance of quantity quality and organization of data through GIS applications through pre-census, census, and post-census periods. Chapter 2 ends by the examinations of the census geographies of the USA and the UK. Chapter 3 firstly gives short background information about statistics in Turkey, and then describes the current statistical activities, necessary steps towards compliance with European Union (EU), NUTS, general population census in Turkey, and small statistical areas. At the end of the chapter, a new Standard Census Geography is proposed for Turkey, and the necessary rules and methodologies are given for further and nation-wide applications. Chapter 4 starts with introducing the case study area, Çankaya, and then gives the reasons of why this area has been selected for this thesis. Then, it informs the reader about the data and the processes of data model design (both conceptual and physical), collection of graphical and non-graphical data, database design, and data manipulations such as conversion, editing, integration etc. Chapter 5 includes basically the real-world application of all of the defined rules and methodologies; and finally gives a prototype of the proposed census geography. Chapter 6 gives the conclusions extracted from the thesis, and some further recommendations for future studies. The flow chart of the thesis is shown in Figure 1.1.

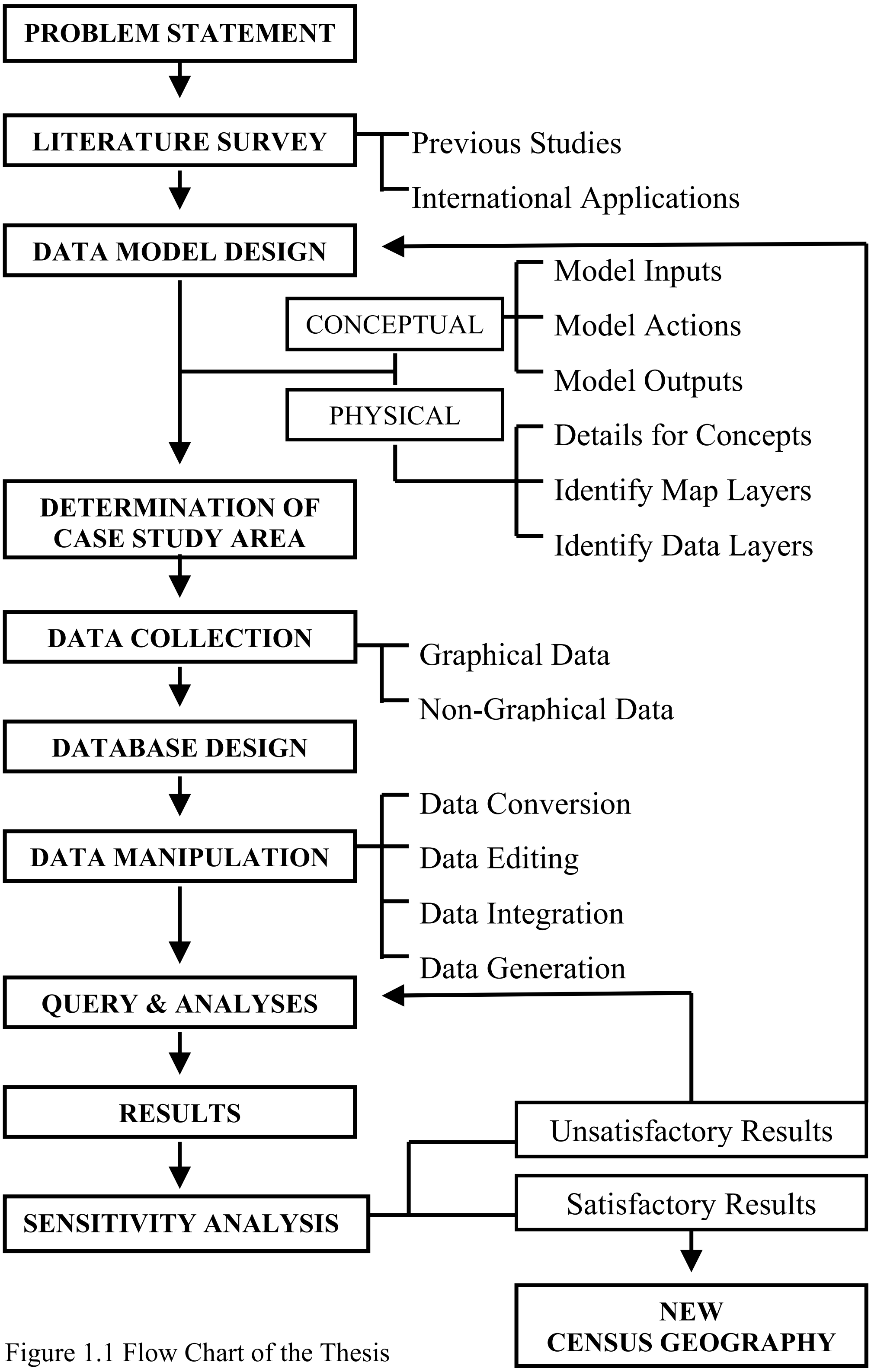


Figure 1.1 Flow Chart of the Thesis

# CHAPTER 2

# GEOGRAPHIC INFORMATION SYSTEMS IN CENSUSES AND INTERNATIONAL APPLICATIONS

## 2.1 Geographic Information Systems in Population Censuses

“Now Population Census is no longer a mere counting of heads; it involves extraction of information, which plays a vital role in determination of many of governmental and private policies” (Singh, 2001). Census data are also mentioned as the raw minerals of a ‘gold mine’ by the Fiji Bureau of Statistics (2000). It is important to have adequate equipment to have the maximum productivity. Data are raw in nature and they have to be processed before being used as information, which is the most valuable concept of today’s world. Ahuja (1999) defines basically three important dimensions about data; quantity, quality and organization.

### 2.1.1. Quantity of Data

Data have to be a short summary of the reality because nobody needs every part of the data. If there is a percentage of data all of them may be used, but if all the data about the study area are possessed one can use only the required data. “In selective collection, it is quite likely that the required data are missed out; so the analogy is to have all the data collected as it is difficult to say at the stage of data collection as to which data will finally be required. But data are costly to collect” (Ahuja, 1999, p.2). Ahuja believes that, in addition to collection cost, storage is also expensive and finding the required data among large quantities of unnecessary data may be quite time consuming. Too many data make it difficult to use in the directions of the study. As result, every unnecessary cost (time, effort, money) decreases the quality of the expected results.

### 2.1.2. Quality of Data

Quality of data is the second dimension. Every researcher sometimes complain about the quality of data handled because of "…the lack of Completeness - Portion of the area of interest, Accuracy - Correctness of data, Precision - Fineness of the scale, Time - Period of validity, or Costing - Money spent etc." (Ahuja, 1999, p.2). In general; the higher the cost, the higher the data quality. For instance, high resolution satellite images give more detailed information and are more expensive than lower resolutions. However, if such detailed data are not necessary, higher data cost may not result higher gains.

### 2.1.3. Organization of Data

The third dimension is the organization of data. Most times, a critical part, organization part is missed through data collection, tabulation of the collected data, analyses and distributions. Whereas, the lack of organization through these processes may cause important loss in the quality of the data and makes the data user's work more difficult. Data organization must be thought as a part of the whole process and must be strengthen by a well data management system.

The importance of census data has been better understood recently and many users have begun to use them more efficiently. Of course the most effective reasons making census data of Turkey very important are the developments in the census system and the increase of the number of questions through questionnaires. Moreover, the analysis and dissemination of census data have become relatively faster in parallel to the availability of powerful computing software and hardware technologies.  However, these data may not represent the real world if they are not georeferenced. 'Georeferenced Data Related Information Technology' will be very important in the near future. Ahuja (1999) says that; " It is the need of every government to ensure that all the statistical data collected by various agencies is well managed in terms of its availability to various users in a most acceptable format where one has not to refer other records, documents and maps."

It is the GIS technology, which has the answer for the State Institute of Statistics of Turkey (SIS) to improve the efficiency and utility of the department's role in Data Management System. There are many definitions of GIS used by different people from different organizations and backgrounds. Generally;

> *"A GIS is a set of principals, methods and tools used to capture, store, transform, analyze, model, simulate and display spatial and non-spatial data. The information resulting from the whole process aims at both understanding and explaining the spatial distribution of phenomena and events, and at supporting planning decision making process. GIS relates a location to an asset or an event"* (Kong, 1999).

Irwin (2001) thinks that, the most important feature of GIS is that it links tabular data to geographic locations. If this feature is handled in census case, for instance, GIS can link tabular data such as labor force, education degrees, average household size, population number etc. to geometric representations of statistical areas. GIS also, as different from a simple computer database program, can be used for representations, queries and analyses based on a spatial perspective. For instance, the answers of questions such as: Which districts are next to District C? How many schools are inside of Province B? What types of businesses are located between the roads A and B? It is impossible to answer these spatial questions using an aspatial database program. GIS is also used to integrate different data types by using the spatial location as the common link. For example, a data user can spatially join fire responsibility zones with data by block groups.

If the extent of questions of the decennial census is broaden and qualified staff relates those data with GIS, the whole concept of census data will be changed forever. GIS will not only make census data more accurate than ever, but also help to generate much more information than the traditional methods and permit graphical representation of the information in space. In this way, data will be more meaningful and decision makers will have accurate information for better decisions.

Indeed, in United Nations (UN) Statistics Division's (2004) opinion, geographic information is as important as physical infrastructure elements such as roads, power lines, sewage, and other public utilities. Economic value of geographic information has been perceived by many sectors such as regional planning, land management, environment/natural resources, health care, transportation, urban systems, marketing studies etc. worldwide.

The geographic information process mainly includes three phases: data acquisition, data management and data dissemination (see Figure 2.1)

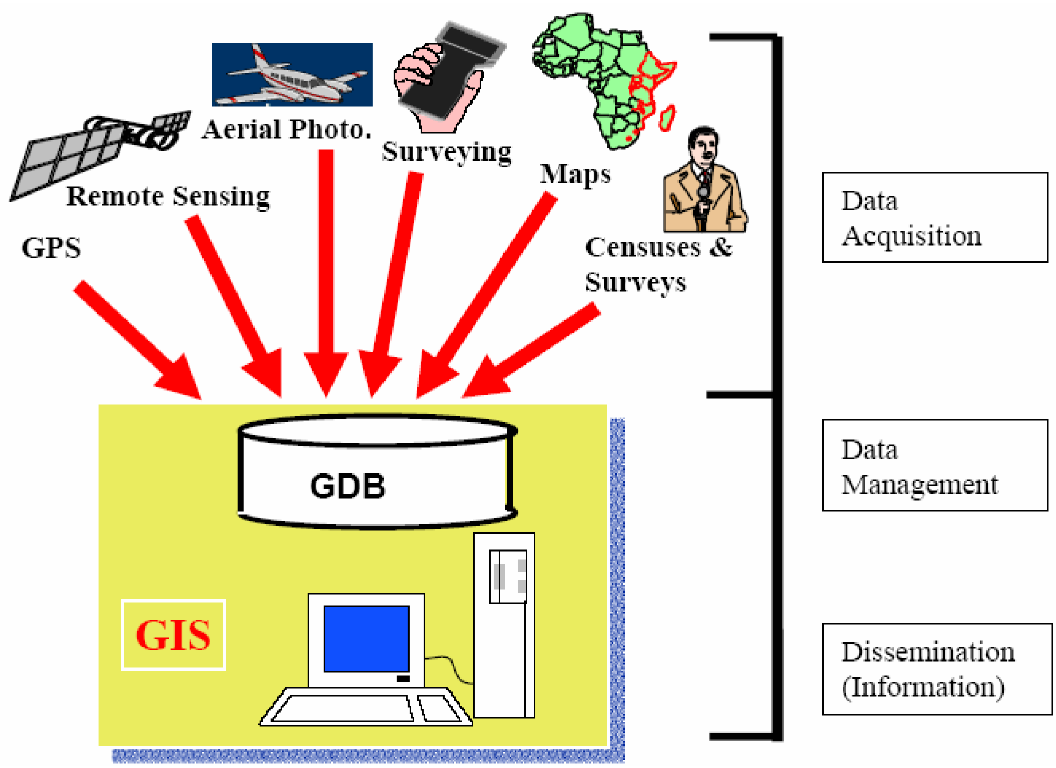


Figure 2.1 Geographic Information process (source: UN Statistics Division, 2004)

It is clear that, "…the GIS technology is necessary to enable more sophisticated use of statistical data and the geographical information, and it offers a better system in terms of standards, methodology, collection, data processing and its dissemination." (UN

Statistics Division, 2004) There is, therefore, a need for the SIS to improve the concept of GIS in pre-census, census and post-census periods of 2010 Population Census.

The pre-census stage maps in Figure 2.2 guarantee consistency and facilitate census operations; then in census stage these maps help to collect more accurate data, to have more information about the study area and to monitor the census process during enumeration; and finally in post-census stage (post-enumeration), GIS and maps make it easier to present, analyze and disseminate census results (United Nations, 2000).

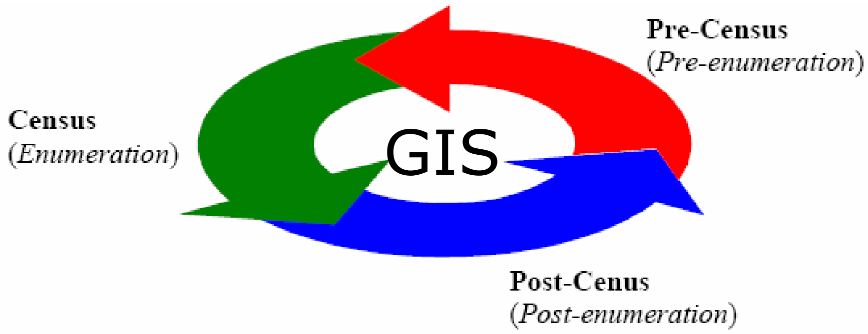


Figure 2.2 Census process (source: UN Statistics Division, 2004)

Census mapping process is very similar to a standard geographic information process. Therefore, it is easy to integrate GIS with census mapping operations in all the stages of the process (see Table 2.1). In this way, census mapping activities go beyond a technical approach, become a united whole with technology and better reflect the population structure in local, regional and nation-wide scopes.

Table 2.1 GIS with census mapping: Stages of integration

| PRE - CENSUS | CENSUS | POST-CENSUS |
|---|---|---|
| GIS<br>GPS<br>Photo / Video<br>Digital Mapping<br>Satellite Imagery | GIS<br>Digital Mapping | GIS<br>Internet |

(source: UN Statistics Division, 2004)

### 2.1.4 Probable Drawbacks of GIS

In addition to the advantages, there are some drawbacks of GIS that is worth considering when implementing the technology and acquiring data. Fiji Bureau of Statistics (2000) gives a list of these drawbacks;

1. System to buy: There will be multiple contracts to be signed for various components of the software and hardware system;

2. Cost-effectiveness: Emphasis should also be placed on the design, creation and maintenance of the database. Data conversion is often an enormous task requiring digital space and time. Database is by far the costliest facet of GIS implementation when acquiring and converting information;

3. Quality Control: Explicit quality control must be exercised throughout the building and maintenance of the GIS database. Without this being conducted, the integrity of data will not be assured;

4. Compatibility with other systems and data: A uniform standard for the computer hardware and software need to be established to allow flexibility in networks and data sharing;

5. Expectations: It is vital to identify what is needed to be accomplished and how it functions before making the commitment to develop a GIS system.

6. Staff Training: Qualified staff is needed to operate the system. More in-house training and refresher course needs to be taken.

### 2.2 International Applications

Turkey needs to reach the international standards in the field of statistics as in many other areas. Thus, examination of the common and already in use international statistical systems would be very helpful while creating a new census geography. United States and United Kingdom examples have been examined because of their strong and confidential statistical systems.

#### 2.2.1 United States Census Geography

Statistical areas of the United States include regions, divisions, urbanized areas (UAs), census county divisions (CCDs), unorganized territories (UTs), census designated places (CDPs), census tracts, block groups (BGs), and census blocks (see Figure 2.3).

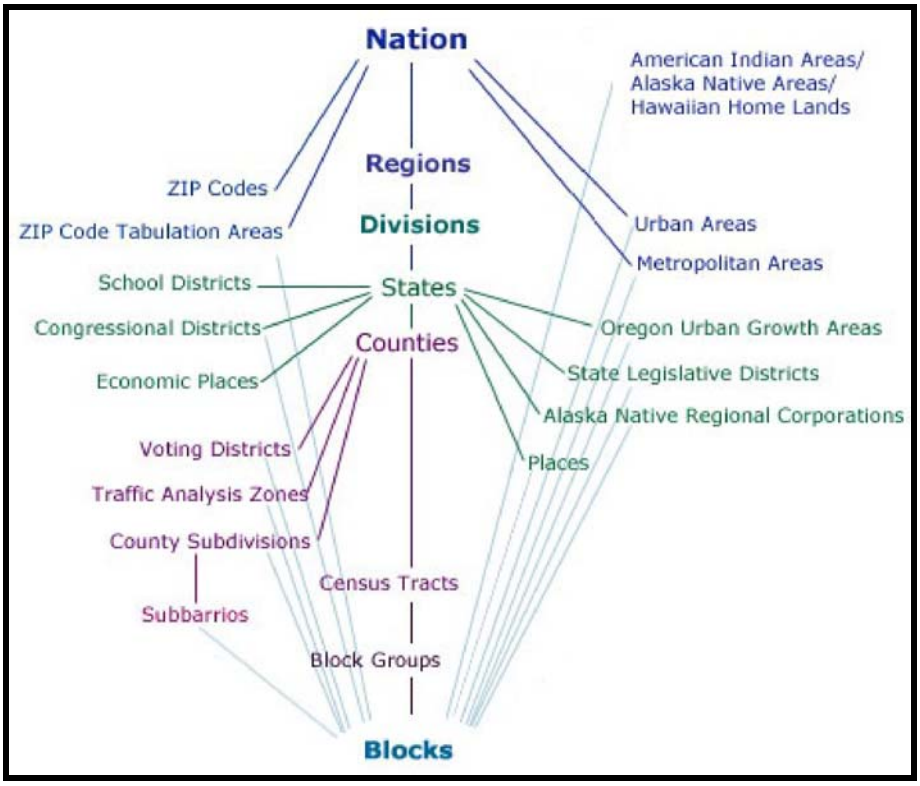


Figure 2.3 Illustration of the legal and statistical entities for decennial census of US
(source: U.S. Census Bureau - Geography Division, 2002)

As it is seen, the census geographic hierarchy is presented as a series of nesting relationships. For instance, a line joining the lower-level entity voting district and the higher-level entity county means that a voting district cannot cross a county boundary in the figure, a line linking block groups and census tracts, means that a block group can not cross a census tract boundary, and so on. There are many statistical entities through the hierarchy, in this way "…the data user community, composed of numerous individuals, businesses, and agencies at all levels of government, each with somewhat different needs, can then select the geographic entity or set of entities that most closely represent their geographic area of interest" (Granda, 2003).

The regions out of the main central line do not necessarily have the lower level entities on the main line. Here the important point is that, all types of regions are related with the most basic unit, the census block. In population censuses, more data are collected for the larger geographic areas, such as the regions and counties; while less data are available for smaller geographic entities such as census tracts and block groups.

The components of the standard census geography in Figure 2.3 are defined in Glossary of United States Census Bureau (2002) as follows:

Region (census geographic); Four groupings of States (Northeast, South, Midwest, and West) established by the Census Bureau in 1942 for the presentation of census data. Each region is subdivided into two or three divisions.

Division (census geographic); A grouping of states within a census geographic region, established by the Census Bureau for the presentation of data. The divisions are subdivisions of the four census regions. The current nine divisions (East North Central, East South Central, Middle Atlantic, Mountain, New England, Pacific, South Atlantic, West North Central, and West South Central) are intended to represent relatively homogeneous areas that are the subdivisions of four census geographic regions.

State; A type of governmental unit that is the primary legal subdivision of the United States.

County; The primary legal division of every state except Alaska and Louisiana. A number of geographic entities are not legally designated as a county, but are recognized by the U.S. Census Bureau as equivalent to a county for data presentation purposes.

Census Tract; A small, relatively permanent statistical subdivision of a county or statistically equivalent entity delineated for data presentation purposes by a local group of census data users or the geographic staff of a regional census center in accordance with U.S. Census Bureau guidelines. Designed to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions at the time they are established. Census tract boundaries are delineated with the intention of being stable over many decades, so they generally follow relatively permanent visible features. However, they may follow governmental unit boundaries and other invisible features in some instances; the boundary of a state or county (or statistically equivalent entity) is always a census tract boundary.

Block Group (BG); A statistical subdivision of a census tract. A BG consists of all tabulation blocks whose numbers begin with the same digit in a census tract. For example, for Census 2000, BG 3 within a census tract includes all blocks numbered from 3000 to 3999. BGs generally contain between 300 and 3,000 people, with an optimum size of 1,500 people. The BG is the lowest-level geographic entity for which the U.S. Census Bureau tabulates sample data from a decennial census.

Census Block; A geographic area bounded by visible and/or invisible features shown on a map prepared by the U.S. Census Bureau. A block is the smallest geographic entity for which the Census Bureau tabulates decennial census data.

### 2.2.2 United Kingdom Census Geography

United Kingdom (UK) Census Geography is not as simple as that of the United States. Office for National Statistics (ONS) (2005a) gives the reasons of this as follows:

- There are many different geographic unit types and their boundaries frequently do not align.
- Boundaries keep changing, and it is said that the UK has more administrative boundary change than the rest of Europe put together. In some years several hundred electoral wards or divisions have been affected, and in the extreme case of 2002 no fewer than 1549 have been changed. Other geographies, especially postcodes, are also liable to frequent revision.
- Area names and codes change too. This is complicated by the fact that different government departments have developed different coding systems for administrative geography, on occasions even using slightly different names for the same areas.

This complex and inconsistent geography has caused many problems when census data are produced and used after census to compare statistics with the other regions. There are of course various solutions to solve this problem. A number of government departments have strong interests in geography but, in many cases, the lead has been taken by the ONS - Geography section. "ONS Geography is responsible for details of a range of products that support the use of geography in the statistical process. Mostly produced by ONS Geography, they have been subdivided into the following categories; Area Names and Codes, Postcode Geography, Health Geography, Administrative Geography, **Census Geography**, Other Geographic Information, Area Classifications, and Digital Boundaries"(ONS, 2005b). Here, the main point to examine is the '*census geography'* of the United Kingdom (U.K.), especially England, in line with the aims of this thesis.

Currently, census in England includes mainly two census geography; collection (Enumeration Districts) and output. Here the output geography also contains incorporation of postcode geography into it and the use of GIS and automated zone design procedures.

#### 2.2.2.1 Enumeration District (ED) design

Martin (2001) says that "The principal consideration of ED design is to create geographical areas which facilitate efficient and accurate distribution and collection of census forms by enumerators, while attempting to equalize enumerator workload." In the previous censuses, EDs were used for both data collection and output geography despite they had been created for only collection purpose. Then it has been understood that, their size and shape are not ideal to be used as Output Areas (OA), so EDs are now used only for data collection. An average size of an ED is approximately 200 households (450 people).

#### 2.2.2.2 Output Areas (OAs) – design principles

The 2001 Census Output Areas are designed specifically for statistical purposes. They are based on data from the 2001 Census and are built from postcode units. According to Martin (2004), output geography design must be based on a GIS database. The first rule is that OAs should have some required minimum and maximum population and household number thresholds to create purpose-specific output areas. The second rule is that, georeferenced census information and postcode geographies should be integrated by creating the OAs from the combinations of the postcode areas. The third rule is that OAs should be standardized to the extent possible according to their population sizes, internal maximum social homogeneity (based on tenure of household and dwelling type), and the shape of areas (especially more irregular geographical shapes). They should have approximately regular shapes and tended to be constrained by obvious boundaries such as major roads. In addition, urban/rural mixes should be avoided where possible. The OAs were required to have a specified minimum size to ensure the confidentiality of data. The minimum OA size is 40 resident households and 100 resident persons but the recommended size was rather larger at 125 households.

#### 2.2.2.3 Output Area Production

Output areas in England are created using fully automated systems with common and consistent criteria across the country. The production system has developed recent years because of the increased power in computing and automatic zoning methods, and also through the availability of co-ordinate referenced and post-coded data. Martin (2004) says that the automatic zoning methods "…make use of the contiguity information available from the GIS containing the unit postcode polygons." Through the procedure, first step is to estimate the number of probable OAs that should fall within a constraining polygon; secondly, according to the given population thresholds, adjacent postcodes are randomly aggregated to form above-threshold OAs; thirdly, some kind of statistical measurements are done to test the results according to design principles defined at the beginning; then the deviations from the target population sizes are measured by the sum of the squared differences between OA populations and the target population size. Here the measurements of social homogeneity and geometric shape are considered separately. Finally, postcode polygons are swapped between adjacent OAs terms of their effects on these statistical measures.

#### 2.2.2.4 Measuring homogeneity

In England system, two of the most effective variables on measuring the homogeneity of the regions are, as mentioned above, dwelling type and tenure; because the structure of the built environment and property ownership patterns may reflect the characteristics of that area. There are, according to the current system, seven dwelling and four tenure categories (see Table 2.2). These two are combined with each other, by equal weights, to determine the almost homogeneous areas. Sometimes 'ethnic groups' data are used while defining the homogeneity but it has little effect on the final decisions.

Table 2.2 Dwelling type and tenure categories in England

| Tenure | Dwelling type |
|---|---|
| Owner-occupied | Detached |
| Rented privately | Semi-detached |
| LA/HA | Terraced |
| Other | Flat |
| | Part-house |
| | Commercial |
| | Non-permanent |

LA/HA: Rented from Local Authority (LA) or Health Authority (HA) (source: Martin, 2004)

**2.2.2.5 Super Output Areas (SOAs)**

SOAs have been created to develop a range of areas that would be of consistent size and whose boundaries would not change. ONS (2005d) mentions that "These have been built from groups of 2001 Census Output Areas (OAs) and will be known as Super Output Areas (SOAs), which have been designed to improve the reporting of small area statistics." For England, to be able to answer a range of potential requirements ONS created 3 layers of SOA, to be used in 2011 census, as described below;

- Lower Layer SOA: Minimum population 1000; mean 1500. Built from groups of OAs (typically 5) The Lower Layer SOAs in England and Wales were generated by a computer program which merged OAs taking into account measures of population size, mutual proximity and social homogeneity.
- Middle Layer SOA: Minimum population 5000; mean 7200. Built from groups of Lower Layer SOAs. The Middle Layer SOAs, were generated via a two-stage process: (1) A draft set was generated by computer, in the same manner as the Lower Layer SOAs. (2) Local authorities and other local agencies were invited to propose changes to the draft boundaries in order to establish SOAs that better met local needs.
- Upper Layer SOA: To be determined; minimum population size 25,000. The nature of Upper Layer SOAs has yet to be determined.

# CHAPTER 3

# STATISTICS IN TURKEY

In this chapter, firstly brief background information is given about the history of statistics in Turkey. Secondly, current statistical activities and the necessary steps towards European Union harmonization are examined. Thirdly, information is given about the decennial population census and the evaluation of census data. Finally, in the light of the first three sections, a new census geography is proposed for the 2010 population census, and its practical application rules and methods are described based on the already examined United States and United Kingdom systems.

## 3.1 Background

The statistical background of Turkey has a relatively long past. In the early years of the Ottoman Empire, it was very important to have statistical information about agricultural areas. For example, between the years of 1326 and 1389, several land population censuses were made. The first general land population census was made between 1512 and 1520, and through the following years it had been repeated every 100 years until the Empire felt down. The first general population census to count the number of inhabitants living in the Empire was taken in 1831 (SIS, 2005a).

19$^{th}$ century is the most important period when the foundations of today's statistical system were established. In the early years, statistical offices were opened in every ministry in Istanbul and in every province and district. In the following century, in 1918, a central statistical department was established and it became responsible for all the statistical activities. Then, in the period of Turkish Republic another centralized department replaced the former one in 1930. Then, until 1962 many titles on economic, social and cultural subjects were published. Nevertheless, Turkey was in a rapid development process and the current system was incapable to answer all the necessities. Thus, in 1962, Prime Ministry State Institute of Statistics (hereafter SIS) was set up. Through the following years, functions of SIS were further reorganized and rearranged by legal decrees taking into consideration Turkey's newly emerging needs and circumstances. Today, in 2005, SIS still retains the main responsibility for the preparation and dissemination of national statistics (SIS, 2005a).

### 3.2. Current Statistical Activities

As it has been mentioned, SIS is the only authorized technical and scientific institute which produces publications to fulfill Turkey's information needs on social, economic, and cultural subjects. “The main function of SIS is to comprehensively determine information needs, collect and compile data, and finally, to present information to its users according to the highest international standards” (SIS, 2005b). SIS has 26 regional offices and 1 branch office in Turkey. The statistics and products generated by SIS are currently used as a guide by governmental institutions and foundations, universities, private organizations, decision makers and researchers.

However, Demir and Toprak (2004, p.4) say that “…as the world entered into a phase, which is commonly described as the process of globalization, statistical offices of many countries come to cope with the challenges of the new demands from decision-makers and researchers.” Therefore, the general situation in Turkey has to be examined to see if there are series of problems to be solved to complete the adaptation process or not.

### 3.3 Steps towards compliance with European Union

The declaration of Turkey as the formal candidate country to the European Union (EU) in December 1999 caused to think about the adoption process to EU in the area of statistics. European Commission has ‘*The Accession Partnership*’ rules, which declare short and long term priorities to the candidate countries. The priorities about statistics for Turkey include the following:

For the short term;

> *"...adopting a strategy for the further development of statistics, in particular demographic and social statistics, regional statistics, business statistics, external trade and agricultural statistics; bring the business register up to EU standards."*

For the long term,;

> *"...adopting EU compatible statistical methodologies and practices, in particular as regards GDP (Gross Domestic Production) estimation, harmonized consumer price indices, short-term indicators, social statistics, business register and balance of payments; aligning macro-economic statistics further with the statistical acquis; and ensuring adequate training of staff and improve the administrative capacity"* (Demir and Toprak, 2004, p.5)

After the Accession Partnership, Turkey began setting itself as a member state of EU; and it was clear that, statistics would play a vital role in attaining the goals. This problem had a major priority among all the others. SIS was aware of that and began the harmonization studies in statistics immediately. Studies about this subject are still underway, and today SIS still has to realize many projects to reach the international standards (mainly European Union), and to reach the aim of improving the statistical system of Turkey. In this framework, a High Level Committee has been established to evaluate the situation and to identify overall and key objectives for the adoption studies of the SIS. Some steps taken on this way are described in the 'Country Paper: Republic Of Turkey', prepared by SIS (2002b). There are mainly eight steps defined in this paper, and one of them is the project of 'Upgrading the Turkish Statistical System'.

SIS has prepared the proposal of this project in November 2001 with the assistance of two EU consultants in order to accomplish the short and long term priorities of

Accession Partnership and National Plan. Its total budget is 15.3 million Euro for 36 months period (2002-2004). The project includes many components to upgrade the Turkish statistical system, but one of them, '*Upgrading of the regional statistical system, introduction of NUTS classification, data collection and dissemination system, and a regional indicator database*' is the focus point of this thesis. This component will be examined, and a new geographic statistical system will be proposed to develop the existing NUTS classifications of Turkey, and to make a contribution to the adoption process.

### 3.3.1 NUTS (Nomenclature Units for Territorial Statistics)

Eurostat (2004) defines NUTS as the regions established by Eurostat to provide a single uniform breakdown of territorial units for the production of regional statistics for the European Union (EU). One of the most important aims of this regulation is to take under control the inevitable change in the administrative boundaries through years in the Member States. They use NUTS regions to minimize the effects of these changes on the availability and comparability of regional statistics. It has the benefit of being well established, considerably stable, hierarchical and well organized to the national statistical regions.

#### 3.3.1.1 NUTS in Turkey

In Turkey, until recently, statistical classifications of regions were made according to administrative divisions (local statistics are still based on administrative boundaries). Turkey, consisting of 81 provinces, partly adapted to the European statistical classification in September 2002. Turkey was divided into 12 NUTS-1 units, 26 NUTS-2 units and 81 NUTS-3 units (see Appendix A). All regional planning efforts are carried and incentives extended on the basis of these NUTS regions. These NUTS regions were created and developed according to the following principles of Eurostat (2004):

1. <u>The NUTS must favour institutional breakdowns:</u>

There may be many different criteria to divide a national territory into smaller regions. These criteria are mainly classified into two groups as normative and analytical. *Normative* regions have some political expressions, and they are mainly based on the tasks allocated to the territorial communities. Necessary population size is defined to fulfill these tasks, and then historical, cultural etc. indicators are evaluated while determining the boundaries. On the other hand, *analytical* regions are based on analytical requirements. Specific regions having similar types of geographical or socio-economic characteristics are grouped together while determining the boundaries of the regions. Member States base their NUTS classifications mainly on normative criteria because of practical reasons of data availability and the implementation of regional policies.

2. <u>The NUTS favours regional units of a general character:</u>

Some areas have the specific characteristics of a certain field of an activity like mining regions, rail traffic regions, farming regions, labour-market regions etc., or certain types of living conditions. These areas may also be used to draw the boundaries of NUTS regions.

3. <u>The NUTS is a three-level hierarchical classification:</u>

Since this is a hierarchical classification, the NUTS subdivides each nation into a whole number of NUTS 1 regions, each of which is in turn subdivided into a whole number of NUTS 2 regions and so on. In addition to these two, it is necessary to establish a third regional level, NUTS 3, which corresponds to a less important administrative structure. However the NUTS Regulation lays down some minimum and maximum population thresholds for the average size of the NUTS regions, in Turkey these population requirements have been partly ignored. For instance, population number of a NUTS 3 region can not exceed 800.000 according to Eurostat's rules; but Istanbul, which is also a NUTS 3 region, has more than 10 million inhabitants. The minimum and maximum thresholds for the average size of the NUTS regions in Turkey are given in Table 3.1.

Table 3.1 Population thresholds of NUTS regions in Turkey

| **Level** | **Minimum** | **Maximum** |
|---|---|---|
| NUTS 1 | 2,507,738 | 10,018,735 |
| NUTS 2 | 871,405 | 10,018,735 |
| NUTS 3 | 93,584 | 10,018,735 |

Source: State Institute of Statistics (According to the results of 2000 Population Census)

The current boundaries of NUTS 1, NUTS 2, and NUTS 3 levels in Turkey are shown in Figure 3.1. below;

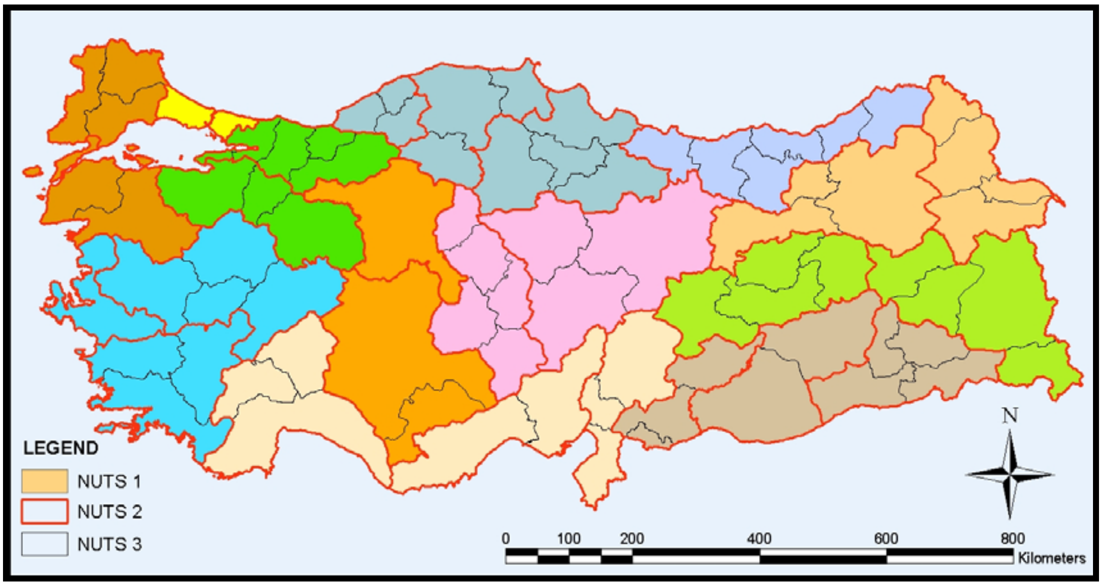


Figure 3.1 Current NUTS 1, NUTS 2, and NUTS 3 regions of Turkey

At a more detailed level, there are ‘Local Administrative Units’ (LAU), which are the next and local levels of the NUTS Regulation. Determining the boundaries of LAUs and creating a new census geography are the main concerns of this thesis; but it is firstly necessary to understand how statistics are produced at present. In Turkey, the main data sources of the statistics are censuses.

### 3.4 Censuses in Turkey

The main census activities maintained by SIS are ‘General Population Census’, ‘General Agriculture Census’, ‘Building Census’, and ‘General Industry and Business Census’. In parallel to the aims of the thesis, only the ‘General Population Census’ will be examined.

#### 3.4.1 General Population Census

The aim of the ‘General Population Census’ in Turkey is to accurately determine the total number of population, and social and economic characteristics of people according to the administrative boundaries. Governmental functions like defense, taxation, and justice require administrative units which are the lower level units of nation. These units are sometimes ‘natural’ or ‘historical’ regions, and they are more or less arbitrary units.

“The success of a census or sample survey depends not only on how well the authorized institution designs the questionnaire, collects the data, and processes the results, but also on how well it links the collected data to geographic areas” (US Census Bureau, 1994a, p.1). As it was mentioned just before, the administrative geography is also used for statistical purposes in Turkey; but it must not be forgotten that; administrative areas are designed mainly for administrative purposes and in most cases not appropriate for statistical purposes.

Comparability of geographic areas from one census to the next is often a major concern for data users, and some users require a stable set of boundaries that permit historical comparisons. This, sometimes, becomes impossible because of the changed, divided, or combined administrative boundaries. For instance, the total number of provinces in Turkey was 63 in 1950, 67 in 1960, 73 in 1990, 80 in 1997; and now, since 1999, this number is 81. That is, 18 new provinces have been added to the original ones in 49 years. The similar changes occur in local levels, too; and these changes make it impossible to generate, analyze and use the time-series data and cause loss of information about related parts of the country.

SIS produces statistics mainly for provinces, and in some cases, for district level. The areas of some provinces are too large for specific regional studies such as environmental impact assessment or market analyses. Thus, it is essential to produce statistics for smaller areas. Furthermore, municipal and metropolitan boundaries are not uniformly regulated across the country, so they may not closely follow urban population boundaries.

It is obvious that geography is a basic element of the census system for organizing and presenting statistical data to the public; but the geography term in the current statistical system needs to be revised. The most effective solution is to create a new **census geography,** which is defined by US Census Bureau (2002, p.9) as "…a collective term referring to the geographic entities used in data collection and tabulation operations, including their relationships to one another." Although the SIS's data tabulations for legal and administrative entities are sufficient to satisfy the needs of many data users, information for these jurisdictions alone does not meet all data needs. Therefore, the SIS must present data for a second geographic category, *small statistical areas*, whose boundaries must be specially defined and have no governmental standing.

### 3.5 Small Statistical Areas

Backer et al. (2002) defend that improving the general quality of the classical systems of official statistics is not the only reason to create a new geographical base for statistics in

Europe. The main reason is to answer the existing and potential importance of information and knowledge systems in production, distribution and consumption of goods and services in the societies. These activities are the main concerns of almost all institutions and organizations taking part through hierarchies of especially public projects. According to them, "…projects contributing to the development of societies are no longer limited to the development of infrastructures within administrative borders, but have been forced to shift their focus to the development of networks whose output patterns are not satisfactorily captured by crude systems of 'large area statistics'." Turkey strongly needs a new system of 'small area statistics', too. This system should not completely substitute the current statistical constitution, rather it must be thought as a complementary to the present system.

A 'system of small area statistics' (SSAS) may be defined both formally and functionally. Formally, it is a system of knowledge based on statistical micro data. It includes statistics, regular and irregular geographical features, and methods. "A SSAS consists of, and is in turn itself a part of and designed to fit into, a constantly changing hierarchical network of processes dedicated to the production and analysis of qualified spatial information" (Backer et al., 2002, p.7). Functionally, SSAS is a collection of processes such as aggregations, benchmarks, and data, spatial, and temporal analysis. It is used "…to improve the results of overriding projects to counter threats and exploit opportunities in view of private and collective efforts to improve the human condition" (Backer et al., 2002, p.7).

The aim of the first 'system of small areas' (SSA) idea was to meet the demand for better information for projects to investigate the opportunities and withstand to present and potential problems. It would also highly improve the comparability of statistics across the EU and make aggregations, dis-aggregations, and re-aggregations of data easier in a network of systems of statistical areas. SSA also raised the quality of data sets by focusing on them and creating higher resolutions when used in combination with methods like sampling, small area estimation etc.

There are standards, guidelines, and criteria for defining, identifying, and delineating the small areas to be used in a population census. Small statistical areas must specify precise criteria for establishing the new component entities. After examining the international applications, basically seven fundamental requirements have been formulated for small statistical areas (SSA); (1) Homogeneity, (2) Functional Integrity, (3) Compactness, (4) Continuity and Consistency, (5) Population Equity with same level of statistical areas, (6) Easy Identification, and (7) Historical Comparability. The explanations of these requirements are given below;

1. SSA should be as homogeneous as possible both socially and economically;

"The homogeneity principle involves combining a group of people, housing units, or business establishments with similar characteristics into a single geographic area (US Census Bureau, 1994b, p.28). This principle is the most important criterion among the others because it minimizes the Ecological Fallacy problem occurring in the current census system. Ecological fallacy is described by Ratcliffe (1999) as a situation that can occur when a researcher or analyst makes an inference about an individual based on aggregate data for a group.

For instance, a researcher looks at the results of a population census and learns that the average annual household income of the residents of 'X district' is $30,000. It is true and accurate to state that the average income for residents of that area is $30,000; there is no problem here. Problem (ecological fallacy) occurs when the researcher states, based on this information, that people living in the area earn about $30,000, however this may not be true for all of the people in the district.

A more focused examination may show, for instance, that the district is actually composed of two housing estates, one of a lower, and the other of a higher socio-economic group. For example, the richer part of the district earns $50,000 while the poorer part earns $10,000. That is, the statement of the census is inadequate because it does not enable people to understand these two different patterns in the study area. Thus, the assumptions made about individuals based on aggregate data are open to the ecological fallacy (see Figure 3.2). Here the important point is to make researchers understand that the process of aggregating or disaggregating data may hide the variations that are not visible at the larger aggregate level.

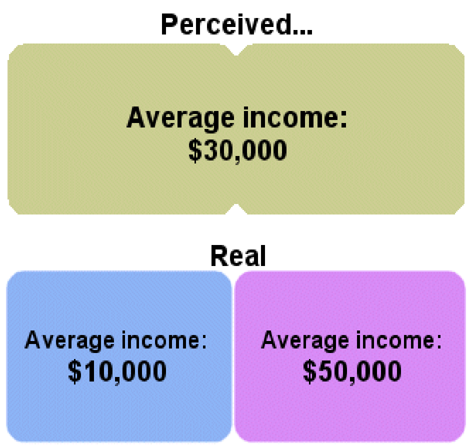


Figure 3.2 Sample for a probable ecological fallacy (source: Ratcliffe, 1999)

There may not be always a distinct boundary which sharply separates two different areas; in opposite there may be kilometers of transition zone that change gradually between the two areas. In such cases, the authorized staff defining the areas must be careful while determining a boundary. They must examine all the possible criteria, which may point to the existence of homogeneous characteristics within an area, before choosing the boundaries. Moreover, through the decision process, they must develop a layout of statistical entities which may involve the combined use of past census data, maps, aerial photography, field observations, and information from local sources.

2. <u>SSA must have functional integrity;</u>

Functional integration principle is also one of the most important priorities while constituting a new census geography unit. "The principle involves the grouping together, into a single statistical area, the people, housing, or business establishments that share a central nucleus along with the surrounding, functionally related entities, such as a large city and its suburbs" (US Census Bureau, 1994b, p.28). Such areas tend to be a whole by both of their geography and functions. The integration inside the statistical unit will be supplied by communication, movement, and interaction of persons, goods, and services. For instance, in practice, a specific urban pattern may fit all the necessary criteria like homogeneity etc. to be named as a small area; but if there is a highway intersecting this area, it also intersects the functional integrity. The life going on there is not a whole and people are not integrated with each other; thus that area should not be defined as a small statistical area.

3. <u>SSA should be as compact as possible;</u>

Compactness of shape is a desirable quality in a statistical entity, particularly for functionally defined ones; and small areas should be as similar by shape as possible to each other. It makes sense in a statistical unit for their peripheries to be approximately equidistant from the centers. Too irregularly shaped areas have less chance to stay invariant about their statistical characteristics than the others. If irregularity is inevitable, it should reflect geographic properties related to the population, housing units or establishments that the area contains. Irregularities may also cause some difficulties for cartographic presentation and may present problems in data analysis as well. Sometimes these irregularities are unavoidable, but the important point is to make the evaluation of all the other criteria carefully and judiciously before final decision making.

4. <u>SSA should eventually have a complete and consistent coverage across the country;</u>

It is important to create a continuous pattern of statistical areas and to define all the areas across the nation without any undefined or empty space. It is also important not to define the same area more than once or not to overlap with another statistical entity. Moreover, a uniform approach should be used to control the identification and delineation of those geographic areas that are numerous and widespread consisting a national consistency.

5. <u>SSA should be comparable in terms of population;</u>

Size of the statistical entities is also an important consideration while creating or changing their boundaries. US Census Bureau (1994b, p.31) say that "…to many data users, *size* refers to the number of people, housing units, or economic establishments within an area rather than the geographic extent of the area." In the context of this thesis, population size guidelines are used as criteria for most types of statistical entities. The size criterion generally determines the optimum and minimum-maximum number of such entities that can be established within a given region. The observation of minimum population size guidelines for statistical entities also helps to ensure confidentiality; while maximum population size helps to prevent excessive population differences between the same types of statistical entities.

In the context of population size, the statistical reliability of the data presented for various geographic units also becomes a significant factor; because many of the SIS's data are collected by sampling method instead of summarizing the complete-count information. The number or percent of sample is very effective on the final data and its confidentiality is measured according to these. In the same way, while subdividing larger areas such as provinces into smaller entities, it is important to define a minimum population number threshold because of the many data items that are based only on sample responses.

6. <u>SSA should be easily identifiable;</u>

While creating a new statistical unit, it is important to make it as simple as possible. US Census Bureau (1994b, p.30) states that "…easy identification and recognition should be key aspects in the wide acceptance of any geographic entity for which the Census

Bureau presents data in its tabulations and publications." The identification of most administrative areas are good examples about this criterion, because boundaries of almost all of them follow easily definable features. Although the boundaries are invisible, local data users are aware of them; but the data users unfamiliar with the territory may need appropriate maps which show the names and locations of most governmental units, and display their boundaries. There are some rules that define the conditions to meet to accept a feature as the boundary of a statistical entity. This is also important for a consistent statistical wholeness, therefore SIS have to use same types of map features as boundaries throughout the Nation. The map features may be visible or non-visible features.

"A visible feature is any cultural or natural element that a person can see on the ground, such as a street, railroad track, above-ground power line, stream, shoreline, fence, sharp ridge, or cliff. A non-visible feature cannot be seen on the ground but is a line that can be located as a line of sight, such as a property line, a short imaginary extension of a dead-end street, or a point-to-point line of sight" (US Census Bureau, 2000, p.8). A legal boundary may also be classified as a non-visible feature. The features are also classified into two groups as standard and non-standard features. A standard feature, which is the most preferable one, is a visible feature that is permanent on its current position and easily identifiable by eye. A non-standard feature is a feature that may not be clearly defined on the area like a ridge, or might be seasonal like an intermittent stream, or might be relatively temporary like a fence of an empty land.

7. SSA should have historical comparability;

It is desirable to maintain comparability of geographic entities from one census or sample survey to the next. Therefore, the boundaries of statistical areas must remain as stable as possible. Cities are living organisms, and of course, numbers and characteristics of the people, homes, and institutions located there change through time.

The point here is to reflect this change to the boundaries of small statistical areas as minimum as possible to preserve the historical comparability of that urban part.

### 3.6 Standard Census Geography Proposal

As mentioned before, SIS should tabulate its data in terms of a geographic entity that does not correspond with combinations of administrative entities; and they must offer a variety of statistical entities in their standard data products. In Figure 3.3, a new hierarchical census geography is proposed to be used in the 2010 Population Census.

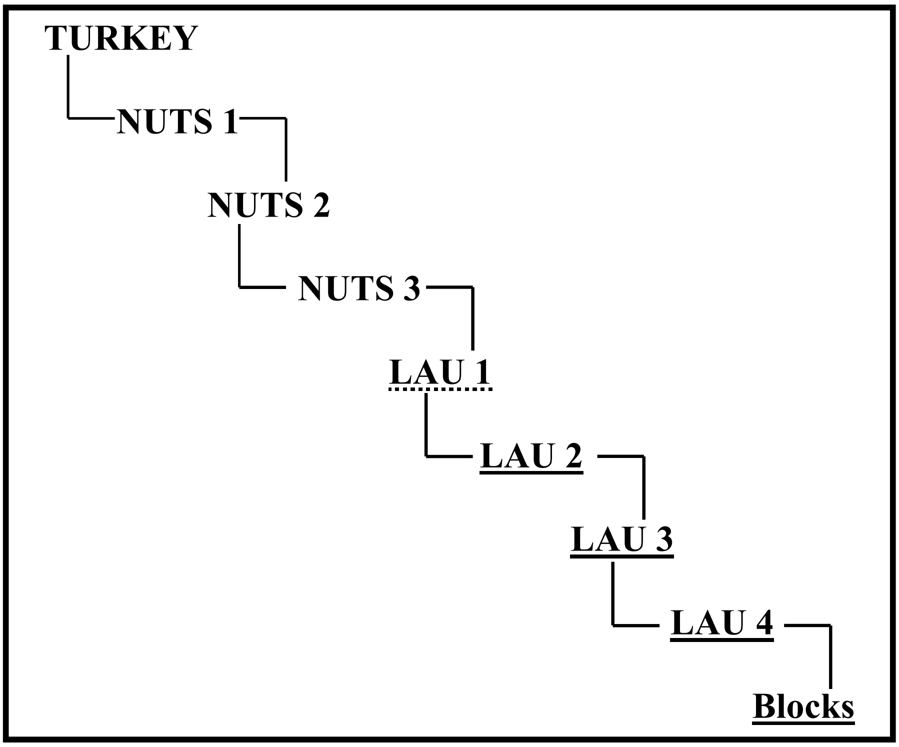


Figure 3.3 Standard Census Geography Hierarchy proposal for Turkey

Geographic hierarchy in the Figure 3.3 shows the geographic entities in a superior/subordinate structure. In this system of relationships among geographic entities, each entity (except the smallest one) is subdivided into lower-order units. NUTS 1, NUTS 2, NUTS 3 regions of Turkey have already been defined by State Planning

Organization (SPO) and SIS. The 'Local Administrative Unit' LAU 1 are assumed as districts throughout the thesis. The main concerns of this thesis are LAU 2, LAU 3 and LAU 4, which are described below in detail. Firstly the smallest units, blocks are defined throughout the thesis, then an upper level, LAU 4, is defined by combining some blocks supplying necessary conditions, then LAU 3 units are defined from LAU 4 units and finally LAU 2 units are created by combining LAU 3 units.

The existing NUTS statistical regions of Turkey were formed according to the European Union criteria, but the lower levels of the statistical hierarchy will be mainly formed according to the United States (US) criteria in the context of this thesis while United Kingdom (UK) criteria are also included in some parts. The higher levels of US could not be used because of two reasons. First, Turkey has already defined the higher level statistical regions; and the second, the administrative hierarchy of US is very different from Turkey. On the other hand, the applications of US for the lower levels of the hierarchy are very suitable to Turkey while UK's lower levels are based on their postcode system, which is very different from the Turkish postcode system. As result, a combination of these two important statistical systems is used, and a new census geography has been proposed. The smallest and basic statistical entities of the hierarchy are blocks.

**Block** → Blocks will be the smallest geographic unit for which SIS will collect and tabulate decennial population census data. Boundaries of blocks are formed by streets, roads, railroads, streams and other bodies of water, other visible and invisible physical and cultural features, and the legal boundaries shown on census maps. Census data for these areas serve as a valuable source for small-area geographic studies, because all the other parts of the hierarchy are mainly based on blocks.

It is a common idea that, census blocks are rectangular or square in shape, in the same size, in a regular pattern etc. although, in fact, the blocks may be very different from each other. Patterns, sizes, and shapes of census blocks vary within and between areas.

The overall pattern of blocks is mainly affected by the factors of topography, the size and spacing of water features, the land survey system, and the extent, age, type, and density of urban and rural development.

US Census Bureau (1994c) describes some rules, which are also relevant for Turkish system, about the shape and minimum size of a potential census block:

- The minimum size of a census block is 2800 $m^2$. There is no maximum size for a census block.
- The width of a census block must be at least 21 m. to prevent extremely narrow slivers being census blocks.
- At least one side of a block has to be a road feature.
- The blocks, which have been planned but do not fit to the existing building and transportation patterns should be redrawn.
- The developed areas which do not have any block structure must have a block structure to achieve a continuous pattern.
- Blocks can intersect with the boundaries of neither buildings nor other blocks.
- Extensions from dead-end streets or pedestrian roads may be used to split oversized polygons into separate blocks.

Many visible and invisible features can be used as census block boundary such as roads, administrative boundaries, drainage systems, feature extensions, rail lines, rail yards, streams, power transmission lines, pipelines, canals, ditches, topographic features, fence lines etc.

**LAU 4** (Block Group) → Block groups (BGs) are the next level above census blocks in the geographic hierarchy. A BG is a combination of census blocks that is a subdivision of a census tract. The BG is proposed as the smallest geographic entity for which SIS will tabulate and publish sample data.

Block Groups are established and can be maintained only for statistical purposes; and they can not be used for any other non-statistical purposes. US Census Bureau (2000) describes some rules, which are also relevant for Turkish system. The general characteristics of block groups are as follows;

- All population and boundary criteria must be met while creating new block groups; and additionally all boundaries must be road features to make the statistical unit accessible except;
  (1) specific legal or land-use areas that are discontinuous,
  (2) discontinuous areas or inaccessible areas that do not meet minimum population size requirements,

- Each census tract must contain a minimum of one BG and may have a maximum of nine BGs.

- BGs must cover the entire land and inland water area of a census tract.

BG boundaries have to follow visible and invisible features, wherever possible; because it is important for BGs to be easily identifiable by users. Legal boundaries are not preferred as BG boundary because of already defined reasons but in some completely urbanized areas, of which boundaries tend to remain unchanged between censuses, may also be used as BG boundary. It must be noted that, all census tract boundaries have to be BG boundaries at the same time. BGs are designed for 100 % count and sample statistics for small geographic areas. There are some population requirements for BGs, which allow some flexibility, to make them meaningful to produce statistics and to provide standardization between different BGs. The population thresholds for the BGs change between 600 and 3,000 inhabitants, with an optimum population of 1,500 inhabitants.

Property areas such as specific residential institutions or military installations containing at least 300 inhabitants, with no optimum or maximum population, are identified as separate block groups. These areas are called special places and they may include large hospitals, prisons; colleges and universities, military installations, hospitals, workers' dormitories, nursing homes, and other institutions with resident population. Separate definition of these special residential facilities not only allows for the collection of data for the residents of such areas but also serves to provide better data about the population outside those facilities.

Population of defined statistical areas may change through time, so some combining or subdividing operations may be necessary. It is necessary to subdivide block groups having population of 3000 or greater. These operations, of course, must retain original block group boundary. This will help data users to aggregate data for the newly created block groups and to compare it with the original single block group area. It is also required to combine a block group into an adjacent block group when the population of a block group is below 600.

**LAU 3** (Census Tract) → A census tract (CT) is a small statistical area, which is relatively permanent subdivision of a quarter to present decennial population census data. CT boundaries have to follow visible and invisible features too, wherever possible; because it is important for CTs to be easily identifiable in the area by data users. CTs are always sub-parts of quarters and their boundaries are quarter boundaries at the same time. CTs should be designed to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions.

Stability through time, so historical comparability, is very important for all of the statistical entities. Ideally, the boundaries of a census tract must also remain the same between censuses to make statistical comparison possible for the same geographic area. Census tracts are established and can be maintained only for statistical purposes; and they can not be used for any other non-statistical purposes. The general characteristics of census tracts are as below;

- All population and boundary criteria must be met while creating new block groups; and additionally all boundaries must be road features to make the statistical unit accessible except;
  (1) specific legal or land-use areas that are discontinuous,
  (2) discontinuous areas or inaccessible areas that do not meet minimum population size requirements,

- A LAU 2 (quarter) boundary must always be a census tract boundary. This criterion takes precedence over all other requirements.

- Census tracts must cover the entire land and inland water area of each province. In coastal waters, territorial seas, and the Great Lakes, it must be created in each quarter a single census tract covering such water bodies to provide for complete census tract coverage.

CTs are designed for 100 % count and sample statistics for small geographic areas. There are some population requirements for CTs, which allows some flexibility, to make them meaningful to produce statistics and to provide standardization between different CTs. The population thresholds for census tracts are proposed to be between 2.000 and 8.000 inhabitants, with an optimum population of 5.000 inhabitants for Turkey. Many areas may be out of the necessary criteria in the following years, and population change over time may create census tracts that are either below the minimum size or above the maximum size. Census tracts with lower population than the minimum threshold must be combined with the closest census tract; and census tracts with higher population than the maximum threshold must be divided to be meaningful and comparable with other CTs.

It is necessary to subdivide census tracts having population of 8000 or greater. These splitting operations, of course, must retain original census tract boundary. This will help data users to aggregate data for the newly created census tracts and to compare it with the original single census tract area. Population of some census tracts may lower through time because of the movement of their original population to somewhere else, for example, to larger cities. It is required to combine such a census tract into an adjacent census tract when its population is below 1500. Comparability, local data needs, and function should play some role in the review and elimination of low population census tracts.

**LAU 2** (Quarter) → A quarter is the smallest administrative unit whose inhabitants need same services, have the same priorities, and neighbor to each other in Turkey. The administrators of a quarter are a headman (called muhtar in Turkey) and a council. Approval of the Governor, decision of Municipality Assembly and opinion of the head official of district are necessary to establish, remove, combine, change and divide a quarter. Headman determines the necessities, improves the quality of life of the quarter by participation of the local people. Furthermore, s/he executes the relationships with municipality and the other public organizations; and notifies information about the quarter. Municipality is the primary authority to serve to the necessities of the quarter (TBMM, 2004).

The LAU 2 regions will be generated by combining the lower level LAU 3 statistical areas. As it has been mentioned, each LAU 3 optimally includes three LAU 4 regions, so each LAU 2 will be generated by combining three LAU 3 regions. According to the proposal of this thesis, the quarters will be redefined, their boundaries will be redrawn according to the already defined criteria, and some of them will be combined while some are divided; but the quarter system will work as it is now, and quarters will again be the sub-administrative units of district municipalities. Of course, all quarter boundaries will be census tract boundaries at the same time.

Quarters, differently from the other statistical entities defined, will serve statistical and administrative purposes. A quarter (LAU 2) must meet the population and boundary feature criteria and comprise a reasonably compact, contiguous land area internally accessible to all points by road; each quarter must contain a minimum of one census tract and may have a maximum of nine census tracts. The population criteria for the quarters change between 10,000 and 20,000 inhabitants, with an optimum (average) population of 15,000 inhabitants. The boundaries of quarters must be defined according to the criteria defined for census tracts.

**LAU 1** (Districts) → Actually, determination of the boundaries of LAU 1 regions is out of purpose of this thesis; because more criteria like political, geographical, historical etc. properties have to be taken into consideration to draw their boundaries. Nevertheless, the districts are assumed as LAU 1 regions to preserve the completeness of the census geography hierarchy, which has 'Turkey' at top and 'block' at the bottom. The districts are the sub-divisions of provinces (NUTS 3 regions) and generated by combining the quarters.

### 3.6.1 Classification of Squatter Settlements

In Turkey, especially large cities include squatter settlements spread over the various parts of the city. Saracgil (1999) defines gecekondu as, "…a process whereby the lower income groups, mainly people from the countryside, incapable of finding a solution to their housing needs within the rules and standards set by society, proceed to solve it illegitimately through their own resources and efforts." Gecekondu areas are irregularly settled, unorganized, unplanned and environments that may be subject to sudden and unexpected changes. Therefore, it does not make much sense to draw unchanging boundaries for these areas. It is clear that these areas will be rehabilitated in either short or long term. Because of these reasons, gecekondu areas are designated as "Temporary Block Groups" in statistical classification proposed in this thesis; and they will be converted to permanent block groups when they do not tend to change over time.

# CHAPTER 4

# CASE STUDY

## 4.1 Location of the Case Study Area

With a population of 758.490 in 2000, **Çankaya**, the case study area of this thesis, is the most populated district of Turkey and located in Ankara. This figure is greater than the populations of 54 provinces of Turkey. The de-facto population of the district reaches 2 million people during daytime. Çankaya has a nation-wide importance because it is the administrative center of the country. It also includes Kızılay (the central business district), national assembly, prime ministry, presidential palace, ministry buildings and foreign embassies.

There are 104 quarters in Çankaya and the total area covered by these quarters is approximately 22,000 hectares. In addition, there are 70,000 commercial offices, 350,000 dwelling units, 9 public hospitals, 166 primary schools, 35 high schools, and 8 universities in the district. 17 percent of the inhabitants live in unauthorized squatter settlements (Çankaya Municipality, 2005).

## 4.2 Why Çankaya has been chosen as the case study area?

Çankaya district has been selected as the case study area mainly because of the diversity it offers in terms of housing, land-use pattern, and socio-economic profile; ease of obtaining data; and its proximity to METU.

### 4.3 Data and Processes

#### 4.3.1 Graphical Raw Data

The next step is to define necessary data to achieve the aims of the study. This study needs both graphical and non-graphical data about the study area. Graphical data are necessary to understand the physical features of the area like urban pattern, topography, transportation, etc. Table 4.1 gives information about the collected graphical raw data.

Table 4.1 Graphical raw data about the study

| Name | Date | Scale | Type | Format | Source | Content |
|---|---|---|---|---|---|---|
| **District Boundaries of Ankara** | 2002 | 1/25000 | Polygon | shapefile (.shp) | AMMOIS | District Name |
| **Quarter Boundaries** | 2002 | 1/25000 | Polygon | shapefile (.shp) | AMMOIS | Quarter Name |
| **Land Use** | 2002 | 1/50000 | Polygon | shapefile (.shp) | AMMOWI | Land Use Type |
| **Transportation Network** | 2000 | 1/25000 | Polyline | shapefile (.shp) | AMMOWI | Name, Road Class, One Ways, |
| **Planned Block Structure** | 2001 | 1/25000 | Polygon | mapinfo file (mif.) | AMMOWI | Block ID |
| **Buildings** | 2002 | 1/1000 | Polygon | shapefile (.shp) | AMMOWI | Building ID |
| **Public Service Locations** | 2002 | 1/25000 | Point | mapinfo file (mif.) | AMMOWI | |
| Hospitals | 2002 | 1/25000 | Point | MapInfo file (mif.) | AMMOWI | Quarter name, Hospital Name |
| Health Centers | 2002 | 1/25000 | Point | MapInfo file (mif.) | AMMOWI | Quarter name, Health Center Name |
| Fire Stations | 2002 | 1/25000 | Point | MapInfo file (mif.) | AMMOWI | Quarter name, Fire Station Name |
| Police Stations | 2002 | 1/25000 | Point | MapInfo file (mif.) | AMMOWI | Quarter name, Police Station Name |

Table 4.1 Graphical raw data about the study (cntd.)

| Name | Date | Scale | Type | Format | Source | Content |
|---|---|---|---|---|---|---|
| **Landmarks** | 2001 | 1/1000 | Polygon | MapInfo file (mif.) | AMMOWI | Landmark Name |
| **Housing Complexes** | 2001 | 1/1000 | Point | MapInfo file (mif.) | AMMOWI | Complex Name |
| **Sample Elevation Points (23757 units)** | 2002 | 1/1000 | Point | shapefile (shp.) | AMMOWI | Point ID, Elevation |

[1] AMMOIS: Ankara Metropolitan Municipality Office of Information Systems
[2] AMMOWI: Ankara Metropolitan Municipality Office of Water and Infrastructure

Each graphical layer described in Table 4.1 has a unique contribution to the study. The Municipality of Greater Ankara has 8 districts in total and 'District Boundaries of Ankara' map is important to see the location of Çankaya district in Ankara and its position according to other districts (see Figure 4.1).

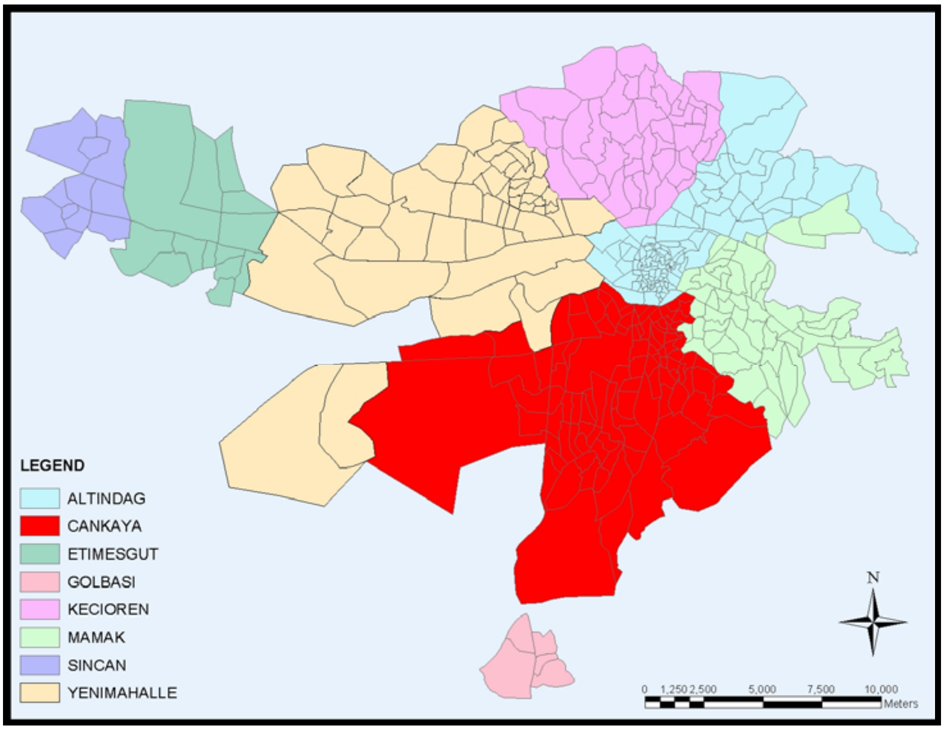


Figure 4.1 Districts of the Municipality of Greater Ankara

Çankaya has 104 quarters, and the 'Quarter Boundaries of Çankaya' layer is probably the most important one because almost all of the necessary non-spatial data are produced at quarter level by the State Institute of Statistics (SIS). Therefore, the boundaries are used to relate the non-spatial data with geographical data. In this way, spatial implications of the collected data are examined and the relatively homogeneous parts are seen better (see Figure 4.2).

'Building' layer is in a 2-D digital representation and there is no information about heights or storey numbers of the buildings. Nevertheless, it is useful for perceiving the building pattern of Çankaya. Building information is used to differentiate regular and irregular settlements from each other, and it is easier to classify the settlements according to their building patterns (see Figure 4.2).

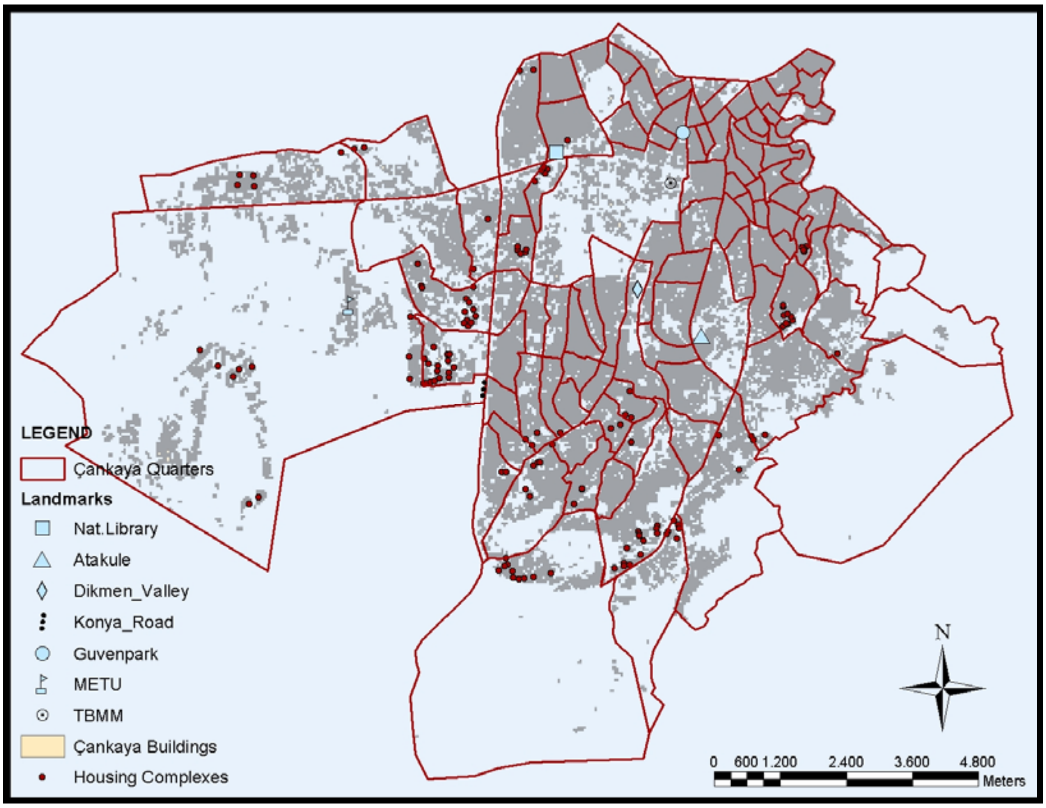


Figure 4.2 Boundaries of Çankaya quarters, building pattern and the locations of housing complexes

Figure 4.2 also includes the layer of 'Location of Housing Complexes'. One of the main rules of creating a new census geography is not to divide the housing complexes, because their dwellers already have the same living conditions and do not need to be divided into separate parts. There are 132 housing complexes defined inside the boundaries of Çankaya. Figure 4.2 also shows that the buildings are denser in the central parts of the district because the East and North sides are generally covered by forests, and the West side contains two university campuses, Middle East Technical University and Bilkent University. It is also known that squatter settlements are located near to the Eastern, Southern and Western boundaries of the district while legally developed areas are concentrated at the central parts.

Çankaya district is classified into 13 types of land-uses (see Figure 4.3).

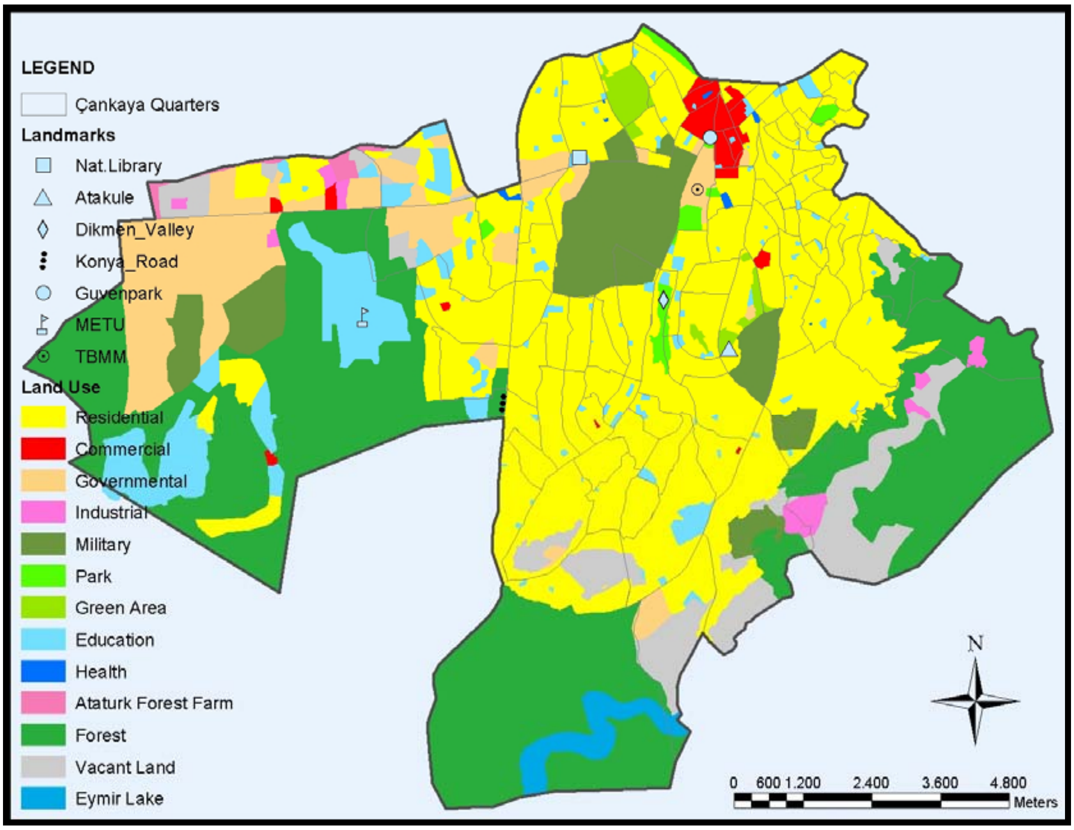


Figure 4.3 Land Use distribution in Çankaya district

The land use classification is mainly based on Corine classification scheme, which is is the EU standard dataset for land cover applications, but it also includes some different classes.The main land uses shown in Figure 4.3 are residential areas (30 %), forests (28 %), governmental uses (8.5 %), military areas (6.5 %), commercial areas (3.16 %) and parks and green areas (1.24 %). The remaining areas are covered by industrial, educational, health care and vacant land uses, together with Ataturk Forest Farm and Eymir Lake (see Table 4.2). 'Land use' layer is necessary to understand where people live and what types of land covers and uses surround them.

Table 4.2 Areal Distribution of land uses

| **Land-Use** | **Area** (ha) | **Percentage** (%) |
|---|---|---|
| Residential | 3899.94 | 30.43 |
| Forest | 3582.89 | 27.95 |
| Vacant | 1770.74 | 13.82 |
| Governmental | 1070.77 | 8.36 |
| Military | 842.29 | 6.57 |
| Education | 618.21 | 4.82 |
| Commercial | 405.59 | 3.16 |
| Green Area | 210.95 | 1.67 |
| Eymir Lake | 158.37 | 1.24 |
| Park | 122.45 | 0.93 |
| Industrial | 89.11 | 0.7 |
| Ataturk Forest Farm | 44.51 | 0.35 |
| Health | 0.11 | 0.001 |
| **TOTAL** | **12815.93** | **100** |

Another layer is the 'Transportation Network of Ankara.' The necessary area was only Çankaya District, therefore, only the roads in the case study area have been taken. The attribute table of roads includes four types of roads: 1st degree (highway), 3rd degree (boulevards), 4th degree (avenues) and 5$^{th}$ degree (streets). This classification was made according to the importance and traffic density of the roads. Transportation network is used to measure accessibility of inhabitants to the public services through urban roads (see Figure 4.4).

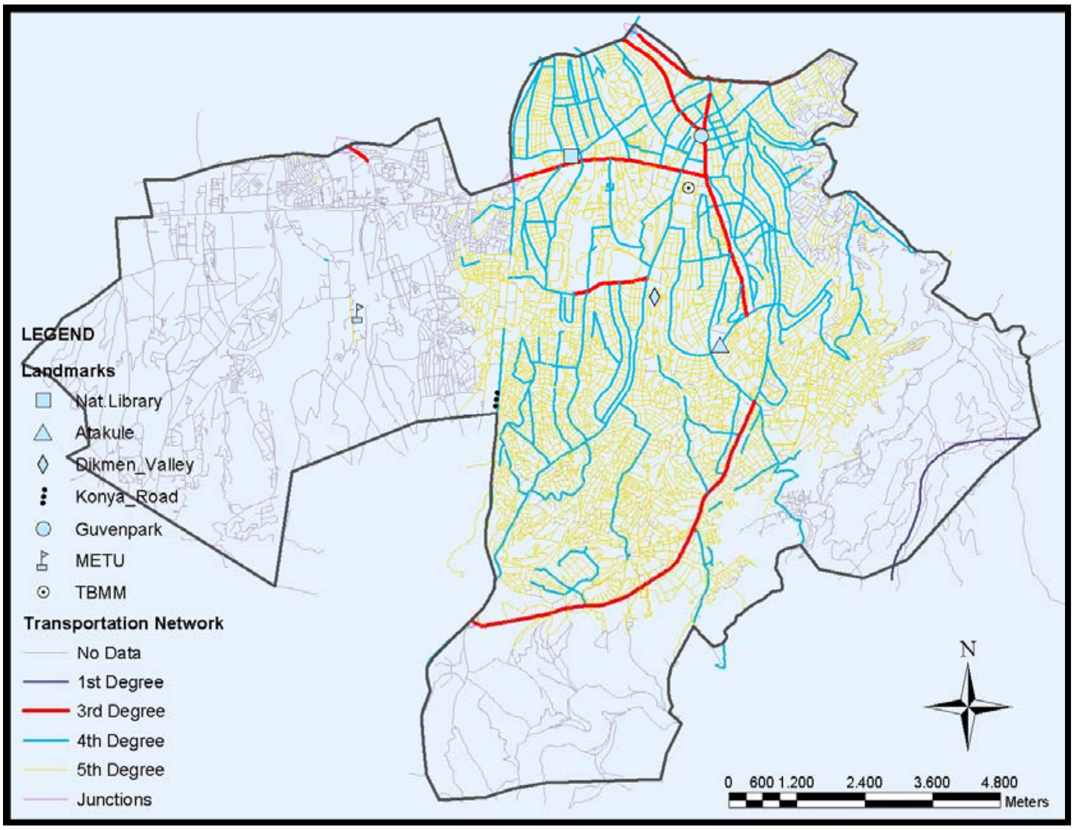

Figure 4.4 Hierarchical transportation network of Çankaya district

'Planned block structure of Çankaya' layer was one of the raw data layers, but it has many drawbacks. First of all, it did not have a complete coverage of the area (see Figure 4.5). In addition, it shows regular block groups all through the district but the actual pattern on ground is very different from the plan. As the figure 4.6 shows, squatter settlements have prevented some areas to develop according to the plan. As a result, there is no harmony between the buildings, transportation system and planned blocks. Because of this lack of harmony, it was impossible to use this layer as it was. Thus, a new block structure, which will be described in Section 4.4, was created for Çankaya using the current planned block structure.

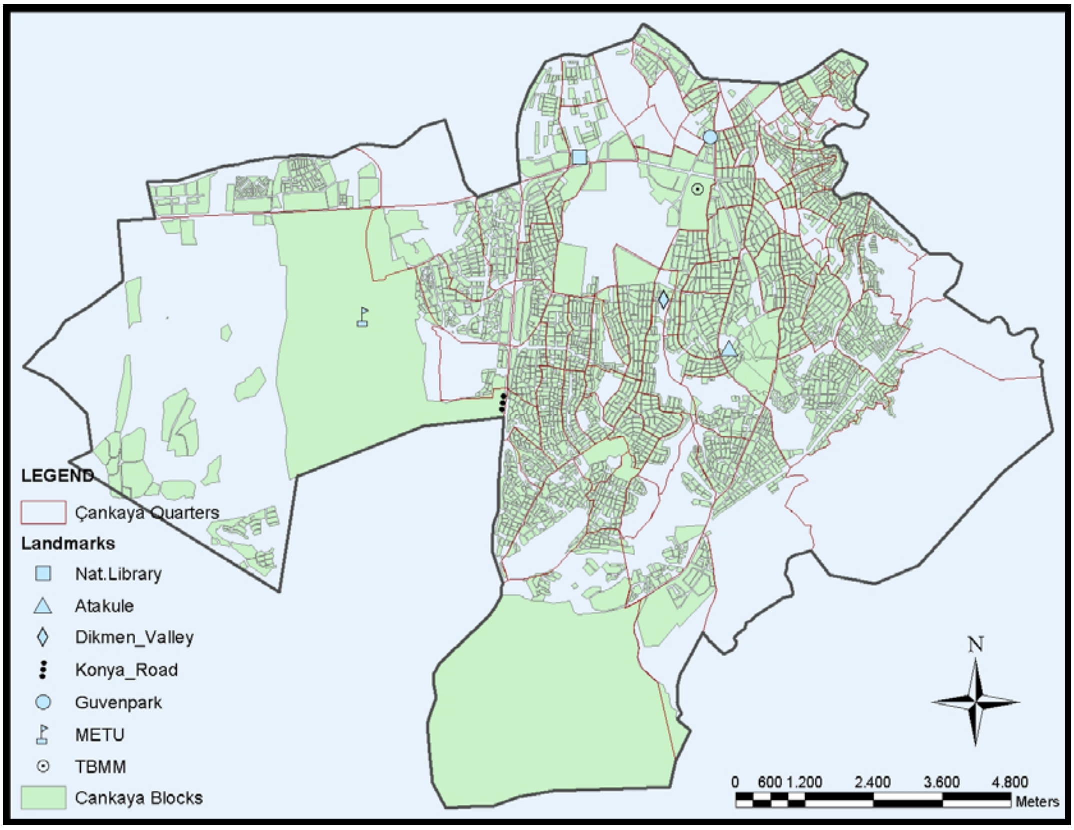


Figure 4.5 Lack of coverage of the planned block structure of Çankaya

Figure 4.6 Disharmony between the planned block structure and the actual pattern

‘Landmark’ layer shows the locations of important and easy-to-find buildings, which is used to perceive the position of settlements according to the important parts of the district. Although there are 80 different landmarks defined in the raw data, they are generally gathered close to each other around the functional center of Çankaya and it is unnecessary to handle landmarks so close to each other in such a small scale study. Therefore only seven of them; National Library, Atakule, Dikmen Valley, Konya Road, Guvenpark, METU and TBMM are used in the representation of layers.

Proximity to public services like hospitals, police stations and fire stations is another important criterion that gives an idea about the quality of life of inhabitants. Therefore, ‘Public Service Locations’ layer will be helpful while querying the safety of people. There are 9 public hospitals, 2 fire stations and 11 police stations located within the boundaries of Çankaya (see Figure 4.7).

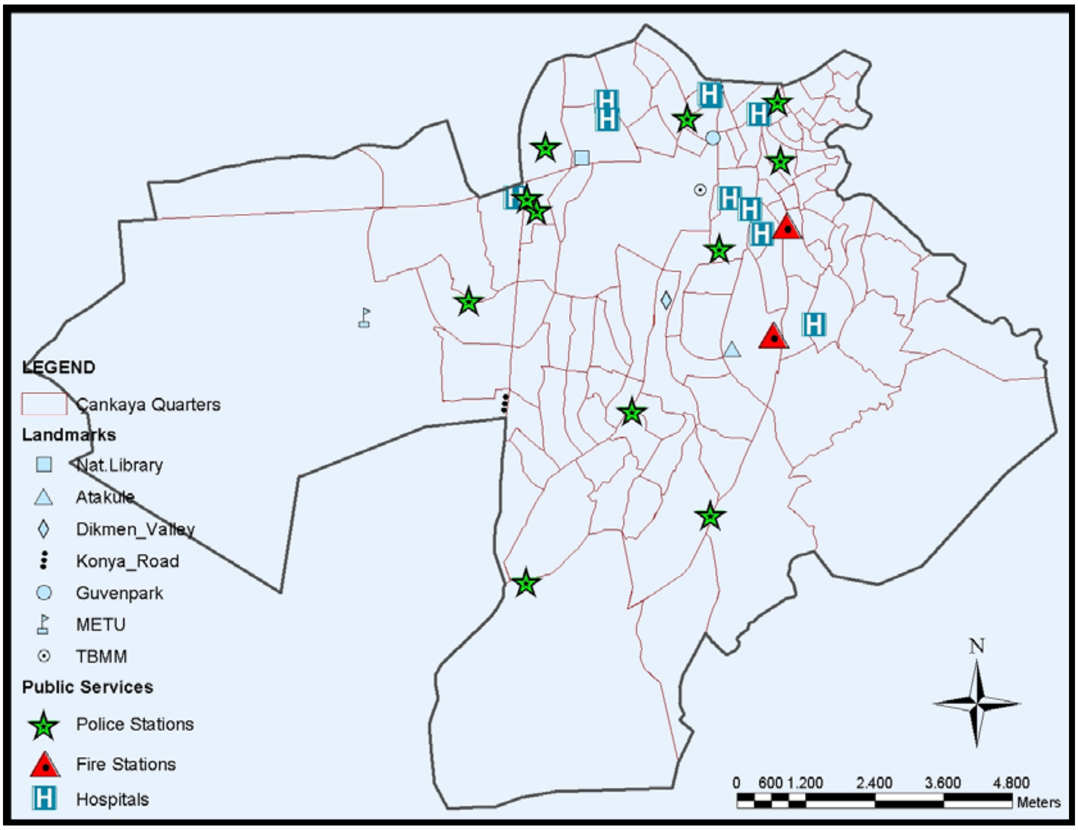


Figure 4.7 Location of public safety services in Çankaya district

'Sample Elevation Points' layer will be used to generate a slope map to understand the topographical features of the area and to help better visualize the district (see Figure 4.8).

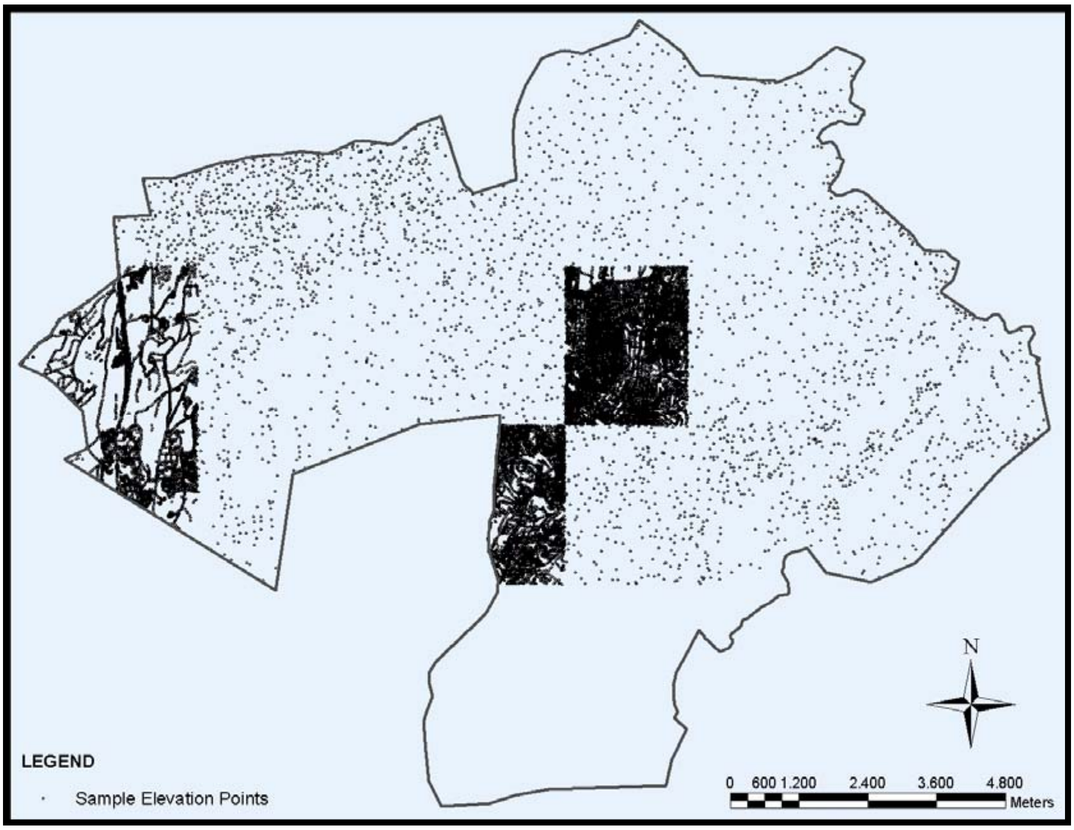


Figure 4.8 Sample elevation points

### 4.3.2. Non-Graphical Raw Data

Non-Graphical data are as important as graphical data, and indispensable for understanding the population characteristics like demography, socio-economic situation, quality of life etc. (see Table 4.3).

Table 4.3 Non-Graphical raw data in the study

| Name | Date | Format | Source | Content |
|---|---|---|---|---|
| **Population of Çankaya Quarters** | 2000 | excel file (.xls) | SIS [1] | Quarter name, Population |
| **Number of housing units in Quarters** | 2002 | excel file (.xls) | SIS | Quarter Name, House Number |
| **Literacy Ratios of the Quarters** | 2000 | excel file (.xls) | SIS | Quarter Name, Literacy Ratio |
| **Graduate Ratios of the Quarters** (%) | 2000 | excel file (.xls) | SIS | Quarter Name, Grad. Ratio |
| University Graduate | " | " | " | " |
| College Graduate | " | " | " | " |
| Middle Sch. Grad. | " | " | " | " |
| Primary Sch. Grad. | " | " | " | " |
| **Unemployment Ratios of the Quarters** | 2000 | excel file (.xls) | SIS | Quarter Name, Unemployment Ratio |
| **Young Age Employment Ratios of the Quarters** | 2000 | excel file (.xls) | SIS | Quarter Name, Young Age Employment Ratio |
| **Quality of Life of the Quarters** | 2000 | excel file (.xls) | SIS | Quarter Name, Quality of Life (Good/Medium/Bad) |
| **Post Code Areas** | 2004 | access file (.dbf) | PTT [2] | Quarter Name, Post Code |

[1] SIS: State Institute of Statistics
[2] PTT: Post Office

The total population of Çankaya is 758,490 according to the Population Census 2000. ‘Population of the Quarters’ is another important layer, because one of the most important criteria while creating new census geography units will be the total number of inhabitants in the quarters. It is essential to note that, the layer could not be used as it was because its scale was less detailed than needed. Nevertheless, it would be useful while estimating the average household sizes and population densities of the blocks. The least populated quarter is ‘Aşağı İmrahor’ with 165 inhabitants, while the most populated quarter is ‘Emek’ with 25.262 inhabitants. These values show that, there are important differences between the quarters in regards to their population. 'Number of Housing units in the Quarters' layer was related with the 'Number of Buildings' layer and used to estimate the number of houses and population of the blocks. The minimum number of houses belongs to 'Aşağı Imrahor' quarter with 54 houses, and the maximum number of houses belongs to 'Emek' quarter, which includes 9680 houses.

Understanding the social structure of quarters is very important in this study. The 'Literacy Ratios' layer based on quarter scale is also an important indicator of the social structure (see Figure 4.9). "The people who know how to read and write with the new Turkish alphabet are accepted as literate. Those who knew how to read but not how to write and those who were able to read and write in the old Turkish alphabet are considered illiterate" (SIS ,2005).

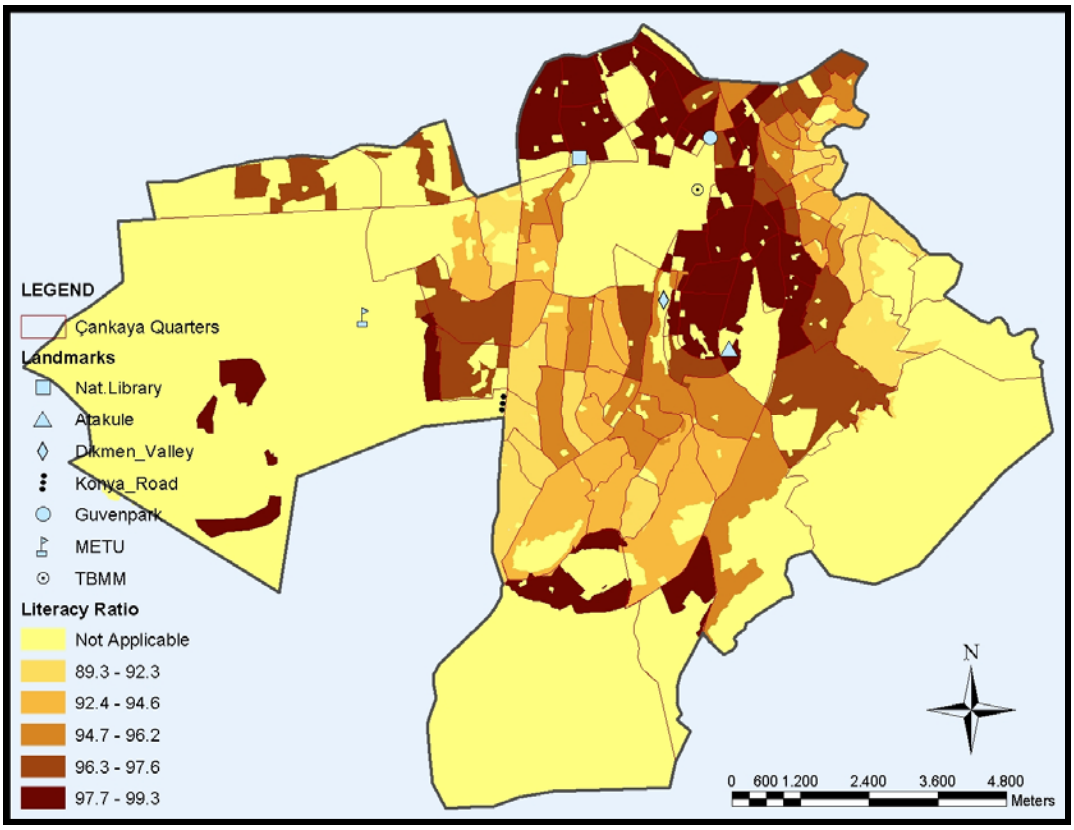


Figure 4.9 Distribution of literacy rate by quantiles in Çankaya District ( % )

The literacy ratio of each quarter was found by dividing the literate people by the total population of the related quarter including only the 6 years of age and over. Figure 4.9 shows that, the ratio of literate inhabitants generally increases from South to North except the most Southern part, and decreases from central locations to Eastern and Western boundaries where low income inhabitants and squatter settlements are concentrated.

'Graduate Ratios' layer indicates the percentage of people graduated from a university, high school, middle school or primary school at the latest. This layer will be used to examine if there is a segregation of people according to their educational levels or not. In this study, only two of these data, ratios of university (see Figure 4.10) and primary school graduates (see Figure 4.11) will be used in the analyses because the others do not show distinguishable differentiation throughout the district.

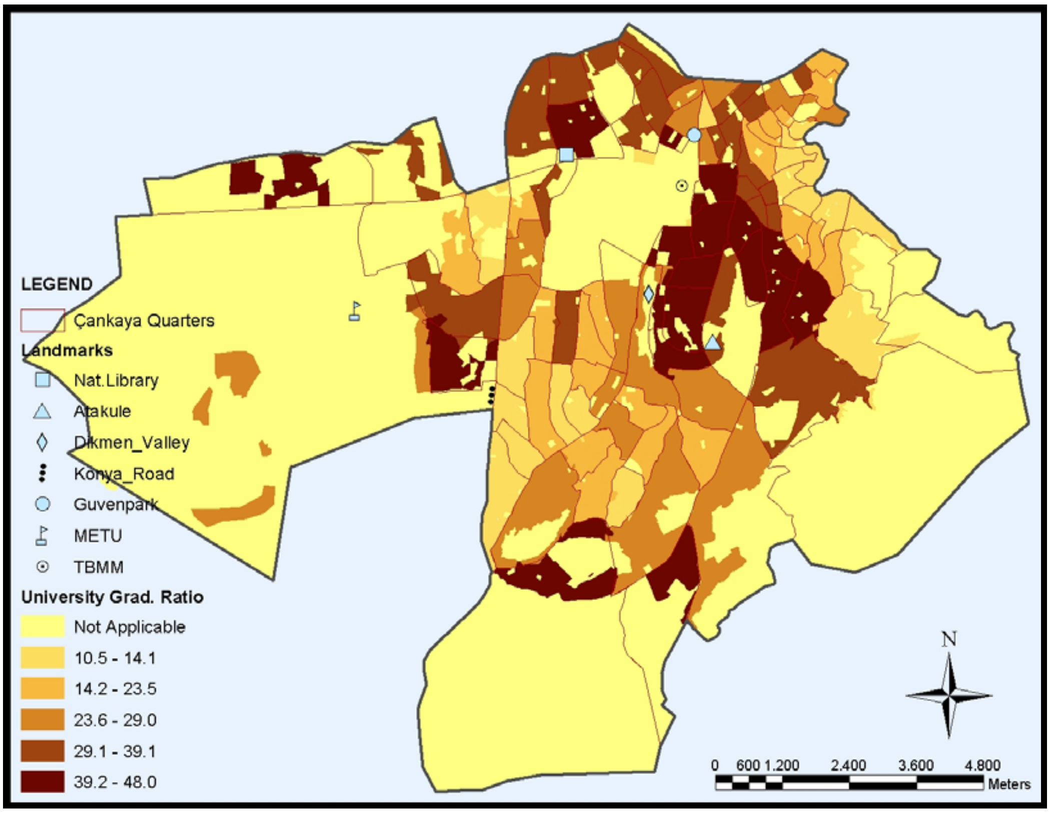


Figure 4.10 Distribution of university graduate rate by quantiles in Çankaya District (%)

University graduate rates of the quarters in Figure 4.10 are found by dividing the number of university graduates by the number of people who graduated from a school, literate, and over 6 years of age. Primary school graduate rates are also found by the same formula. Figure 4.10 shows similarities with the thematic map describing the literacy ratios. The ratio of university graduates also increases towards the city center in Northern part while decreases towards the Southwestern and the most Eastern parts.

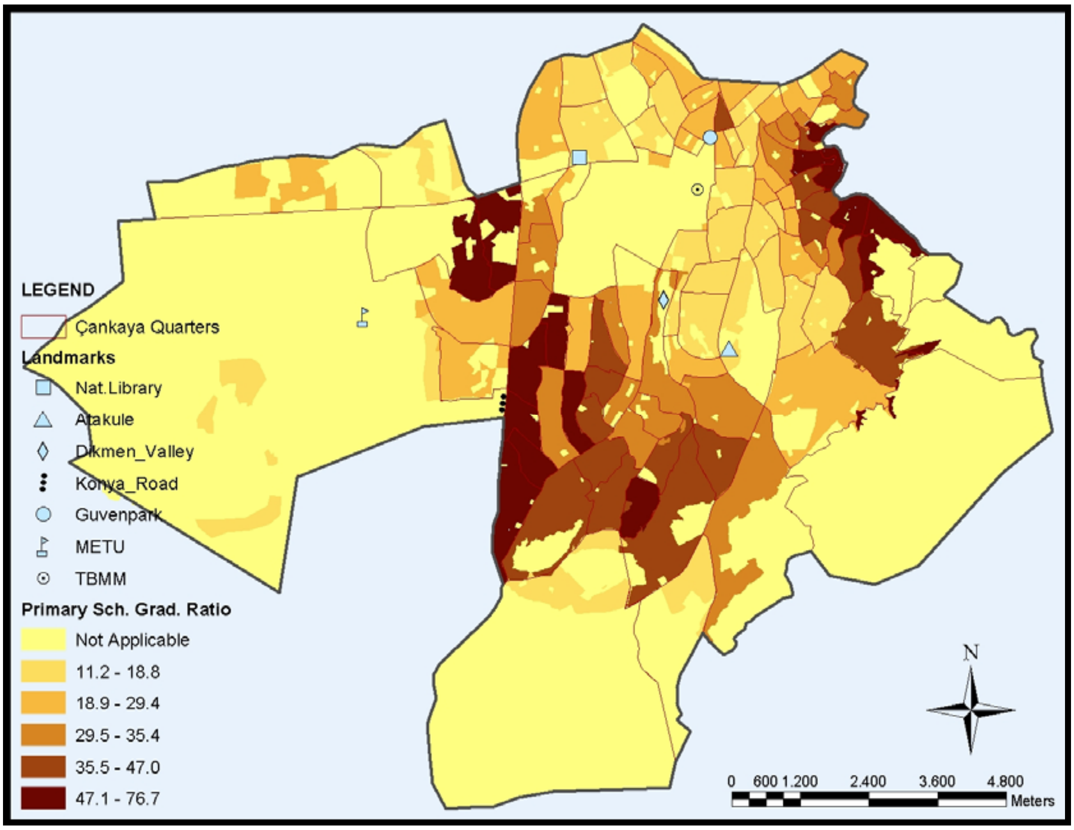


Figure 4.11 Distribution of primary school graduate rate by quantiles in Çankaya District (%)

Figure 4.11 indicates that the ratio of primary school graduates shows a pattern contrary to that of university graduates. The proportion of primary school graduate inhabitants increases going out from geographic center of district to the Southwestern and Northeastern parts of the district.

‘Unemployment Ratios in the Quarters’ layer (see Figure 4.12) and ‘Young Age Employment Ratios’ layer (see Figure 4.13) will also be useful to understand the differentiations among different parts; because they are both considerable signs of low quality of life. According to SIS (2005); “The unemployed comprise all persons 12 years of age and over who were not employed (either worked for profit payment in kind or family gain at any job even for one hour or with any job attachment) during the reference period.” These people have also taken specific steps to work in a job for last six months and have no obstacle to start working in 15 days. Then, the unemployment ratio is the ratio of unemployed persons to the labor force, which comprise the population 12 years of age and over. Employees in 12-19 years of age are accepted as young age employees, and the young age employment rates of the quarters are found by dividing their numbers by the labor force.

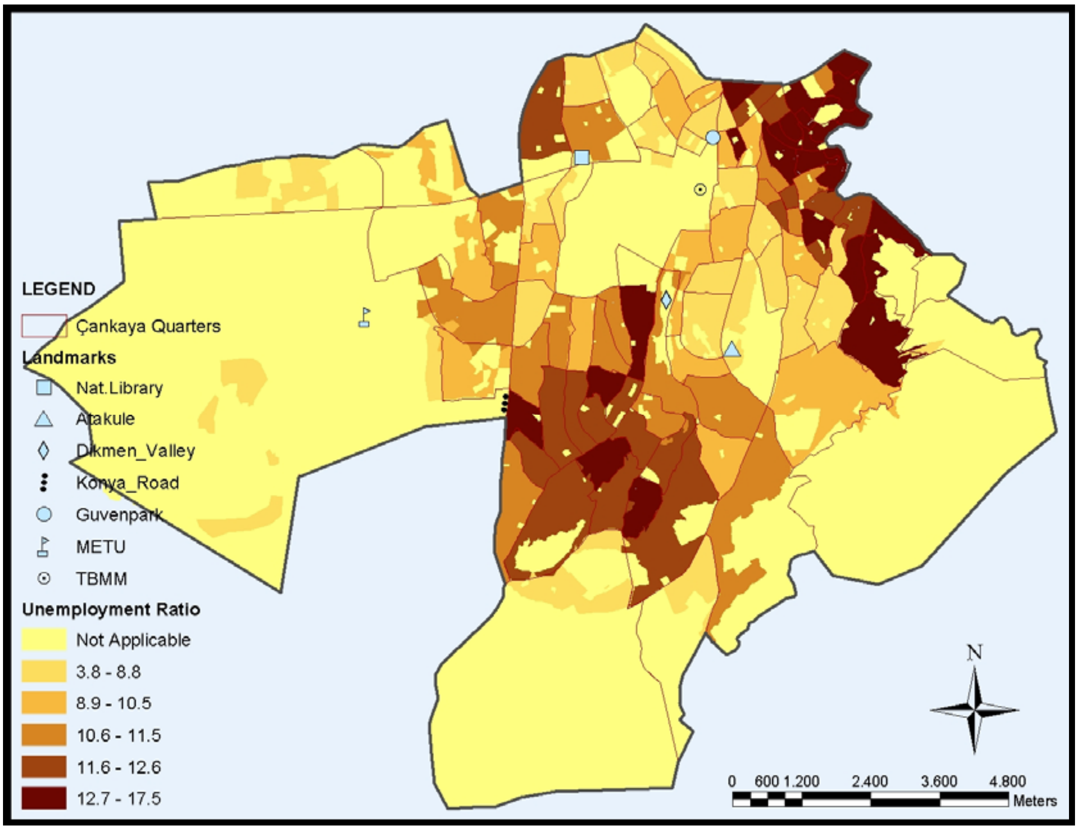


Figure 4.12 Distribution of unemployment rate by quantiles in Çankaya District (%)

Figure 4.12 indicates that, especially the Northeastern parts, where squatter settlements are concentrated, suffer from unemployment. The Southwestern parts also show higher unemployment ratios than the other parts.

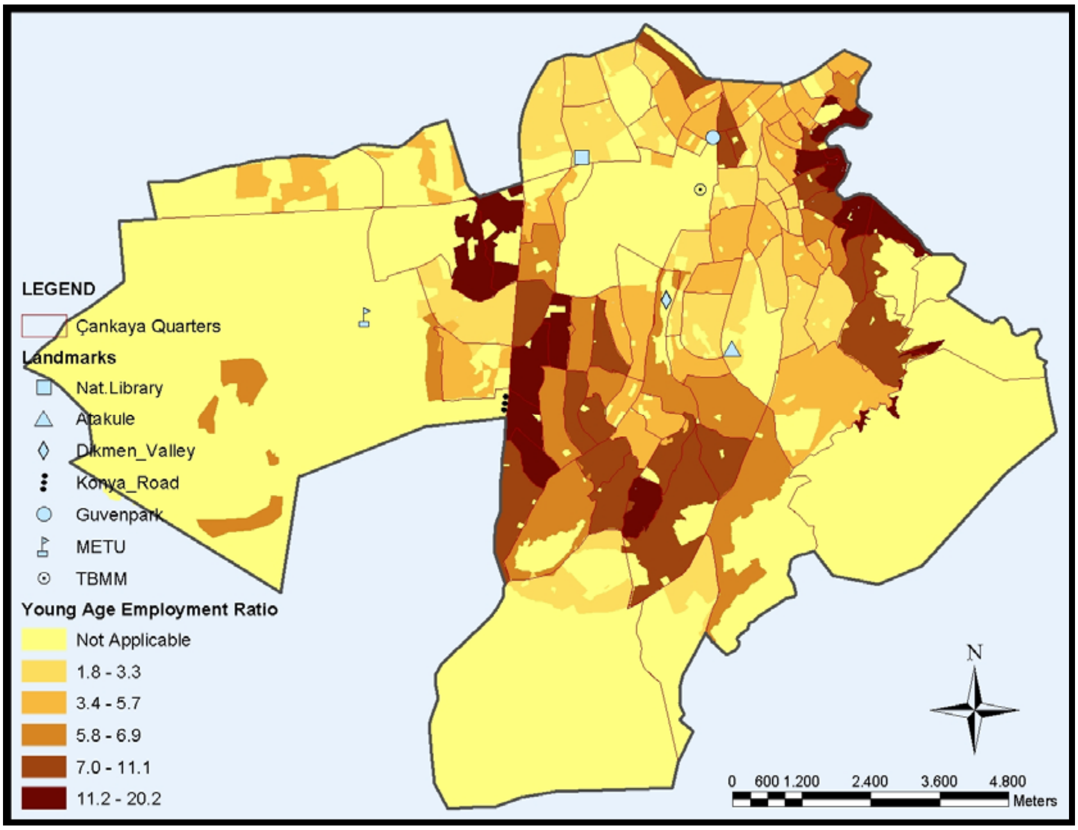


Figure 4.13 Distribution of young age employment rate by quantiles in Çankaya (%)

According to the Figure 4.13, numbers of young age employees, ages between 12 and 19, are higher in squatter settlements that are concentrated in the Northeastern and Southwestern parts of Çankaya than the other parts.

Previous 'Quality of Life of the Quarters' layer is a result of a very simple study, which was executed by the Municipality of Greater Ankara according to the opinions of the municipality employees and without following any theoretical technique. It was generated by making observations in each quarter and finally by assigning one of ‘good’, ‘medium’ or ‘bad’ terms to each quarter. Nevertheless, it is beneficial when forming an opinion about quality of life distribution in Çankaya (see Figure 4.14).

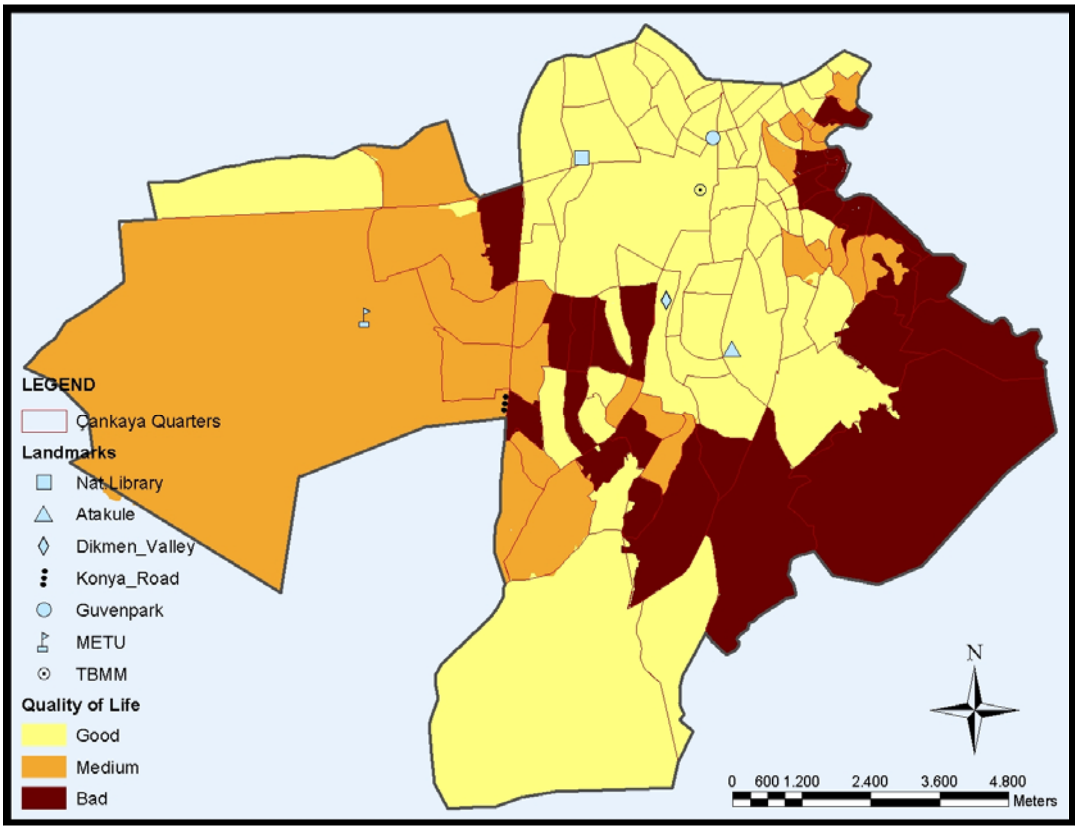

Figure 4.14 Previous approach to quality of life distribution in Çankaya District

According to the thematic map above, the Southeastern parts of Çankaya have bad living conditions, the Western part have a medium quality of life, and the most qualified living standards are located on Northern and Southern parts.

‘Postal Code Areas’ layer is not directly related with the aims of this thesis. Nevertheless, it is helpful to see the uselessness of current postal code system to be used as a base for the new census geography units as in UK case.

### 4.3.3. Generated Data

In addition to graphical and non-graphical raw data, a few more data have been generated from the raw data (see Table 4.4). These data are as important as the raw ones.

Table 4.4. Generated Graphical and Non-Graphical Data about the study

| Name | Date | Scale | Type | Format | Content |
|---|---|---|---|---|---|
| **Population Density of the Quarters** | 2004 | --- | --- | excel file (.xls) | Quarter Name, Pop. Density (# of people per hectare) |
| **Average Household Size of the Quarters** | 2004 | --- | --- | excel file (.xls) | Quarter Name, Household Size |
| **1-3-5-8-10 minute drive zones of the public services (police-fire stations and hospitals)** | 2000 | 1/25000 | Polygon | shapefile (.shp) | --- |
| **Kriging Interpolation Map** | 2001 | 1/1000 | Raster | .aux | --- |
| **Slope Map** | 2001 | 1/1000 | Raster | .aux | --- |
| **TIN** | 2001 | 1/1000 | Raster | .aux | --- |

'Population Density of the Quarters' layer has been generated by dividing the population by the total area of settled residential areas for each quarter, which can be automatically calculated by ArcMap software. This layer is important to see the locations of highly dense areas (see Figure 4.15). According to Figure 4.15; central parts, Northeastern and North regions of Çankaya have the highest population density while the Southern settlements have the least, because of the large vacant areas and low height buildings concentrated in the Southern parts.

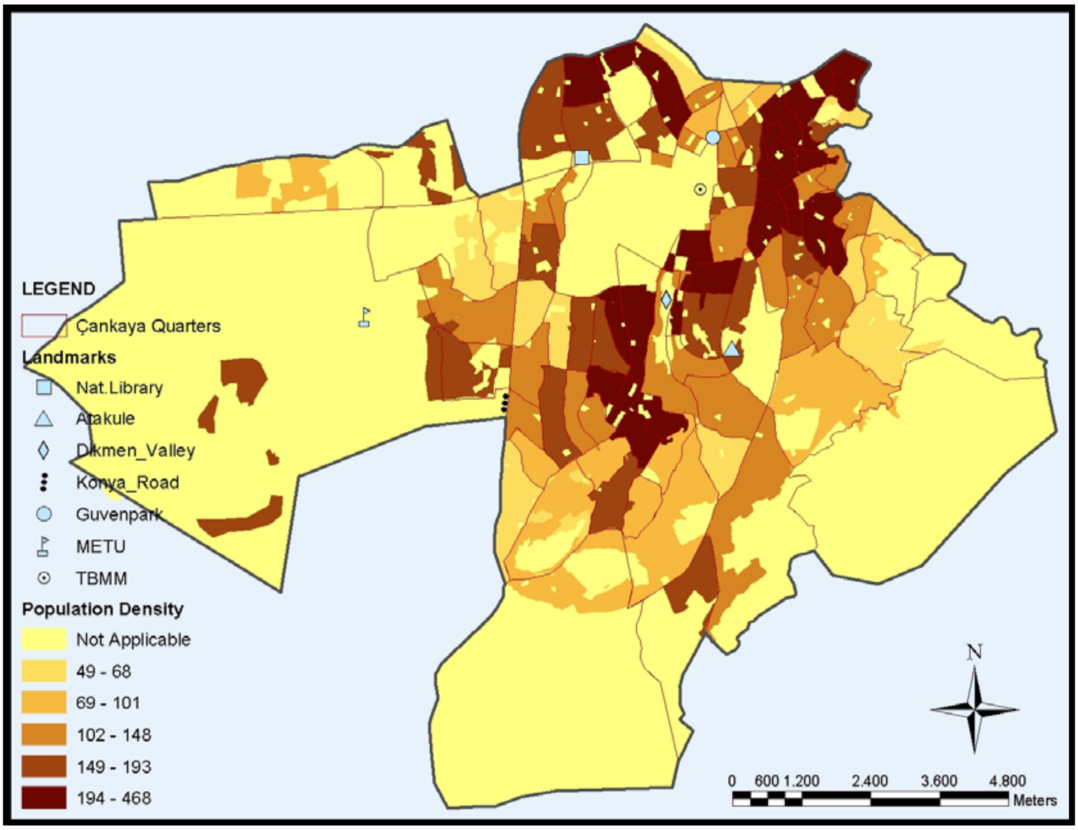


Figure 4.15 Distribution of population density by quantiles in Çankaya (person/hectare)

SIS (2005) defines a household as "...a person or group of people with or without a family relationship who live in the same house or in the same part of a house, and share their meals, earnings and expenditures who take part in the management of the household and who render services to the household." 'Average Household Size of the Quarters' layer has been derived from the division of population by the number of housing units for each quarter. Having small household size is seen as an important demographic indicator of quality of life, and helps to segregate housing clusters from each other (see Figure 4.16).

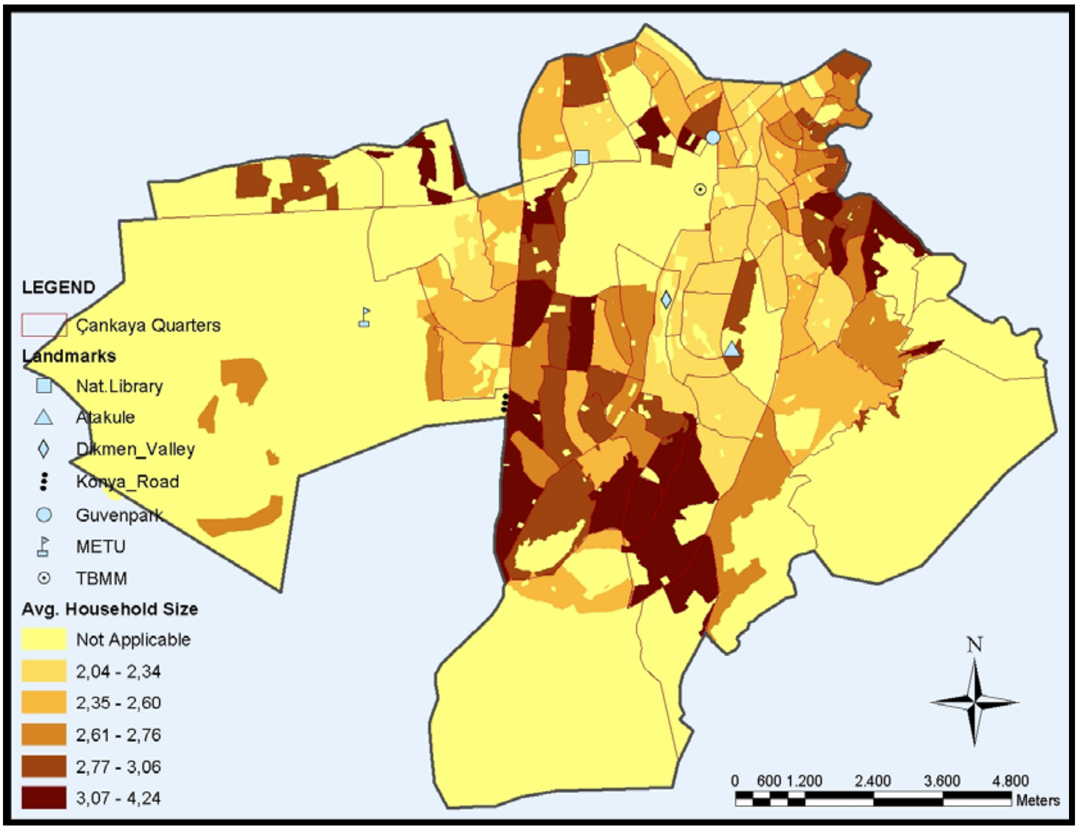


Figure 4.16 Distribution of average household size by quantiles in Çankaya District

As can be seen in the figure, all possible average household sizes are dispersed heterogeneously throughout the district. It is observed that the largest household sizes are concentrated in Southern and South-Western parts and a little in the Eastern part.

'1, 3, 5, 8 and 10 minute driving zones of public services' data were found by connecting the locations of public services layer to transportation network layer by the help of Network Analyst extension of the ArcView 3.2 software. A network is a vector data structure made of *nodes* and *links*. Nodes are the intersections between the links. Each link may also have a number of shape points that define the shape of the feature but do not connect to other links. Network analysis is the mapping of paths of vehicles, people, water, or other objects through map features, often including shortest-route and shortest-time calculations.

The ArcView 3.2 Network Analyst (AVNA) extension helps to solve a variety of problems using geographic networks such as finding the most efficient travel route, generating travel directions, finding the closest facility, or defining service areas based on travel time. In the context of this thesis, **Find Service Area** function of AVNA was used to determine how much area that a particular public facility can serve within a given time or cost frame.

There is a Master of Science thesis called 'Measurement and Evaluation of Fire Service Acessibility by Using GIS: A Case Study on Çankaya District, Ankara', which was written by Kıvanç Ertuğay in 2003. In this study, the necessary analyses were performed and accessibility of fire stations within 5, 10, 15, 20 minutes for heavy, normal and light traffic conditions were found. Based on this thesis, proximity analyses were performed with some modifications.

It was adequate for this study to perform analyses according to normal time traffic conditions. The time thresholds were smaller (1, 3, 5, 8, and 10); because this is a block based and a detailed study. In addition, service zones of hospitals and police stations were also included in the analyses, differing from the thesis of Ertuğay. The maximum speeds of vehicles in normal traffic conditions are determined as in Table 4.5.

Table 4.5 Assigned speeds for the roads in network analysis

| **RoadTypes** | **Speed (V)** |
|---|---|
| **Highways** (1$^{st}$) | 110 |
| **Boulevards** (3$^{rd}$) | 80 |
| **Avenues** (4$^{th}$) | 65 |
| **Streets** (5$^{th}$) | 40 |

Figure 4.17, Figure 4.18 and Figure 4.19 show the results of network analyses performed according to the assigned speeds in Table 4.5.

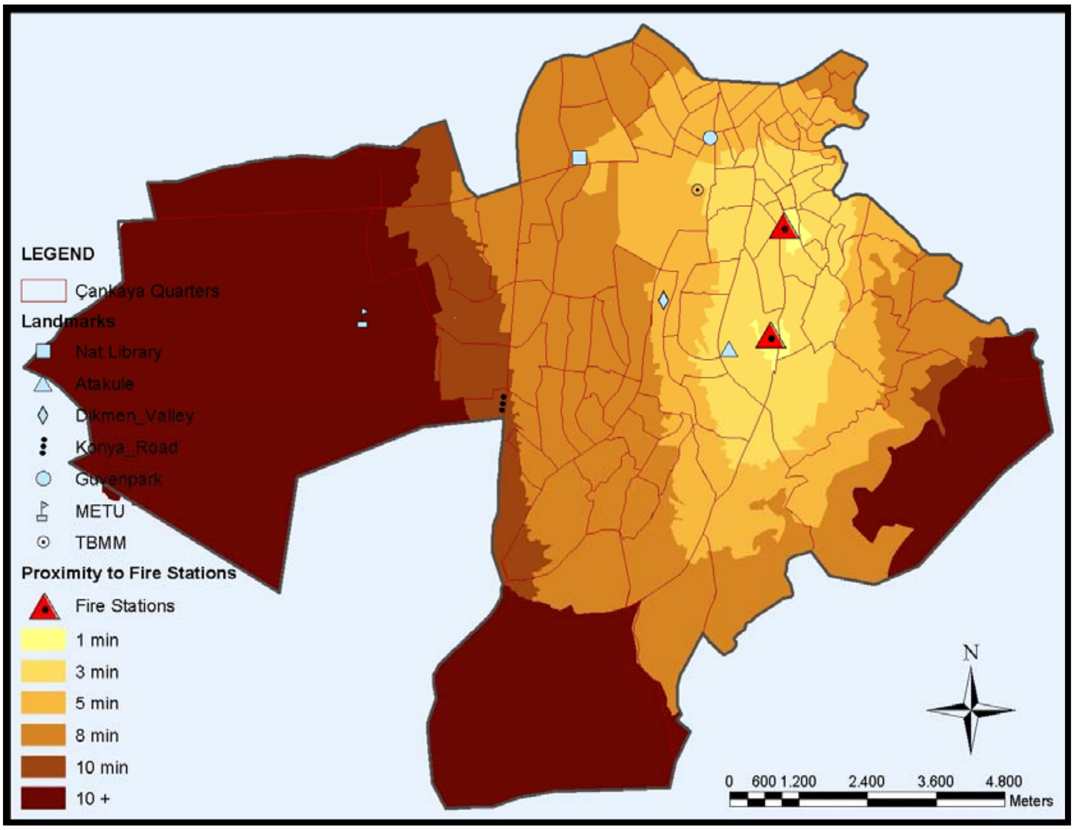


Figure 4.17 1, 3, 5, 8 and 10 minute service areas of fire stations in Çankaya

As the figure above shows, while the central and Eastern parts of the district are very close to fire stations, especially forests and university campus areas at the Southern, Southeastern and Western parts have the minimum accessibility to fire stations.

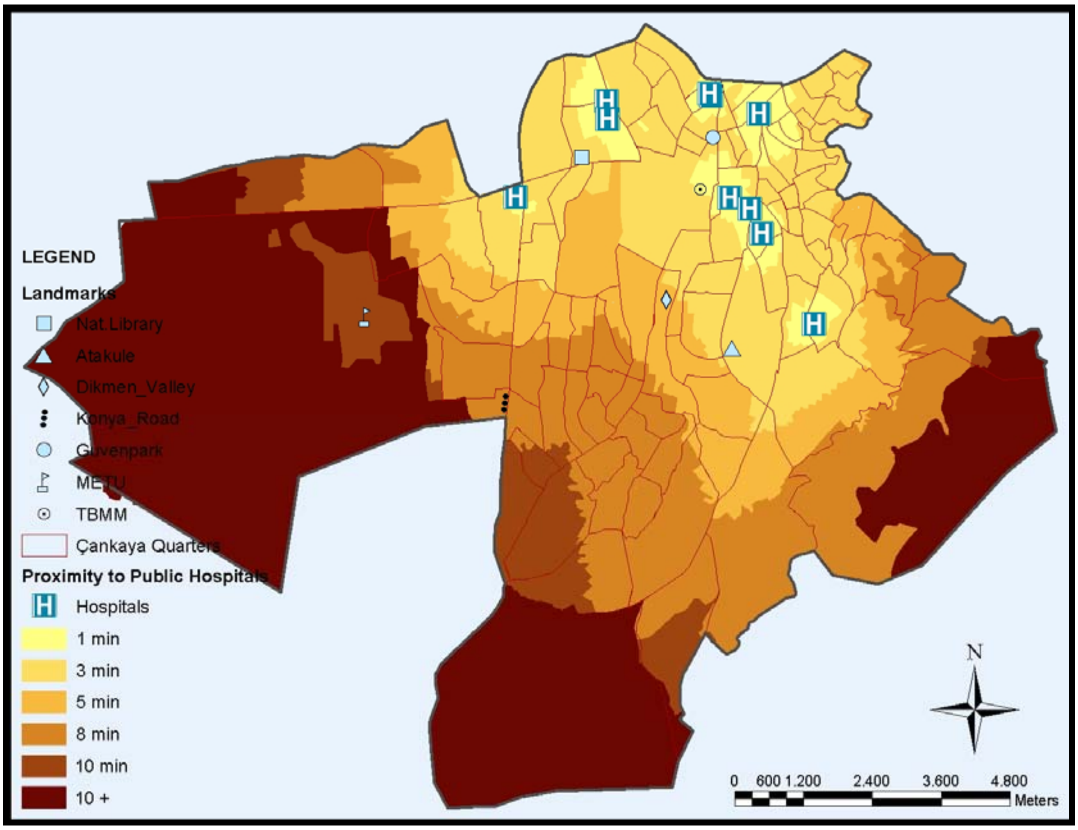

Figure 4.18 1, 3, 5, 8 and 10 minute service areas of hospitals in Çankaya

Figure 4.18 indicates that university campuses and forests have less accessibility to hospitals than any other part of the district, as it is in the fire station accessibility map. Central North and Northeastern parts are the most fortunate areas about health facilities.

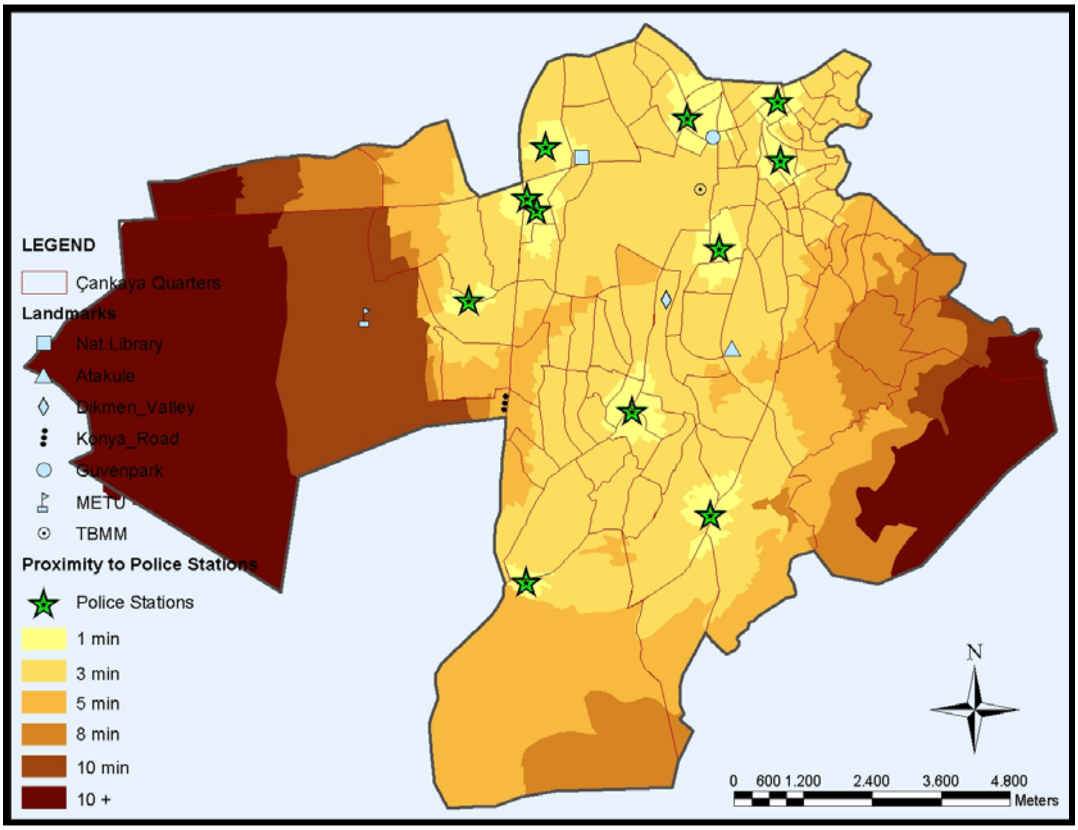


Figure 4.19 1, 3, 5, 8 and 10 minute service areas of police stations in Çankaya

According to the Figure 4.19, Çankaya seems to be in good condition in regards to accessibility to police stations in general except the most Eastern and most Western parts, where there is no need to be protected by police forces.

'Kriging Interpolation Map', shows the estimated height values of the whole district (see Figure 4.20), and a 'TIN' (Triangulated Irregular Network) layer has been generated from the 'Sample Elevation Points' layer. These data will be used to visualize the topographical characteristics of the district.

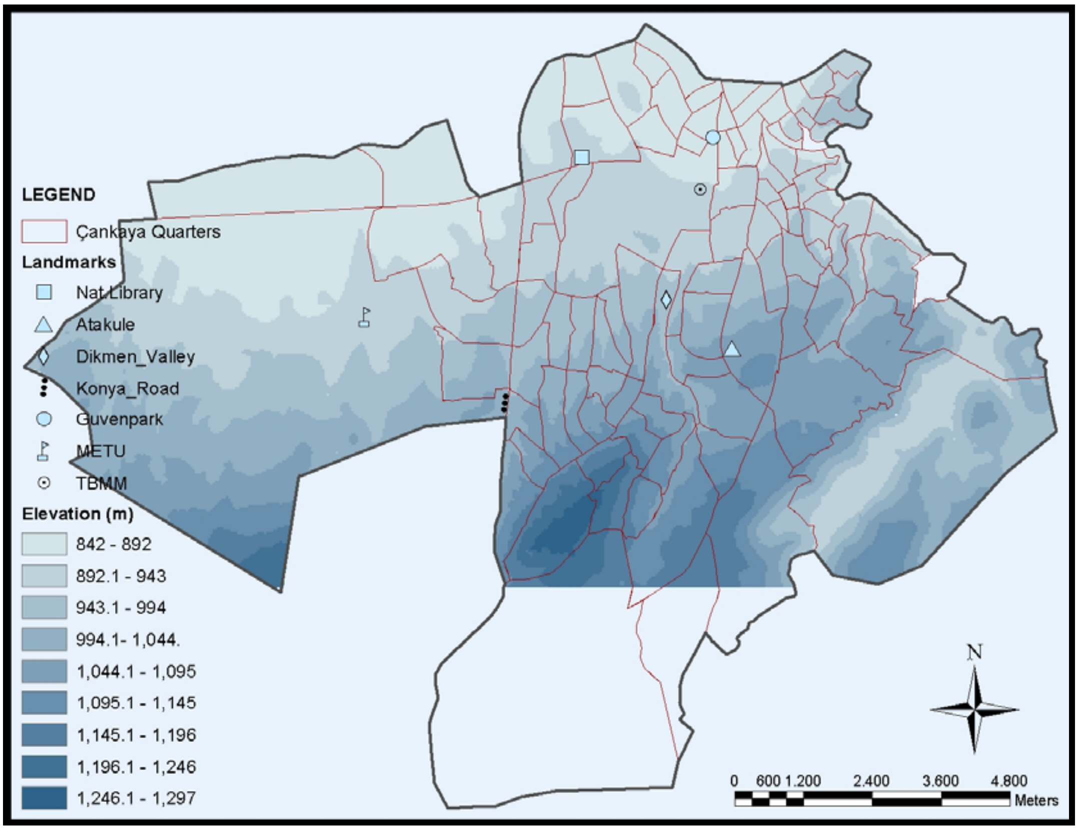


Figure 4.20 Kriging interpolation map derived from the sample elevation points

The map shown in Figure 4.20 was generated in ArcMap software by the help of 'Spatial Analyst' toolbar, 'Interpolate to Raster' command. After activating this command, firstly input points were defined, the field which have elevation values was selected from the attribute table of points, ordinary kriging method was selected rather than universal because of its more accurate results and semi-variogram model was selected as spherical. Search radius type is defined as variable, then in the search radius settings, number of points was defined as 12, and output cell size of the output raster accepted as the software proposed, 40 m. Then, the location of output was shown and the kriging interpolation map was produced.

The next step was to transform the kriging map to 3-Dimensional in order to better express the topographical characteristics of Çankaya and to make it more understandable. The software packages used for this purpose were ArcMap and ArcScene, modules of ArcGIS software. Firstly, in ArcMap the 3D Analyst toolbar was activated and the 'Convert rasters to TIN' option was selected. Secondly, the kriging map was entered as the input raster, automatically proposed Z tolerance accepted and TIN of Çankaya was produced after defining the output location. In the third step, the TIN was opened in ArcScene and the map came to the screen as a scene. Then, in the 'scene properties' command, vertical exaggeration of the scene was increased 7 times and the 3-D representation was generated (see Figure 4.21).

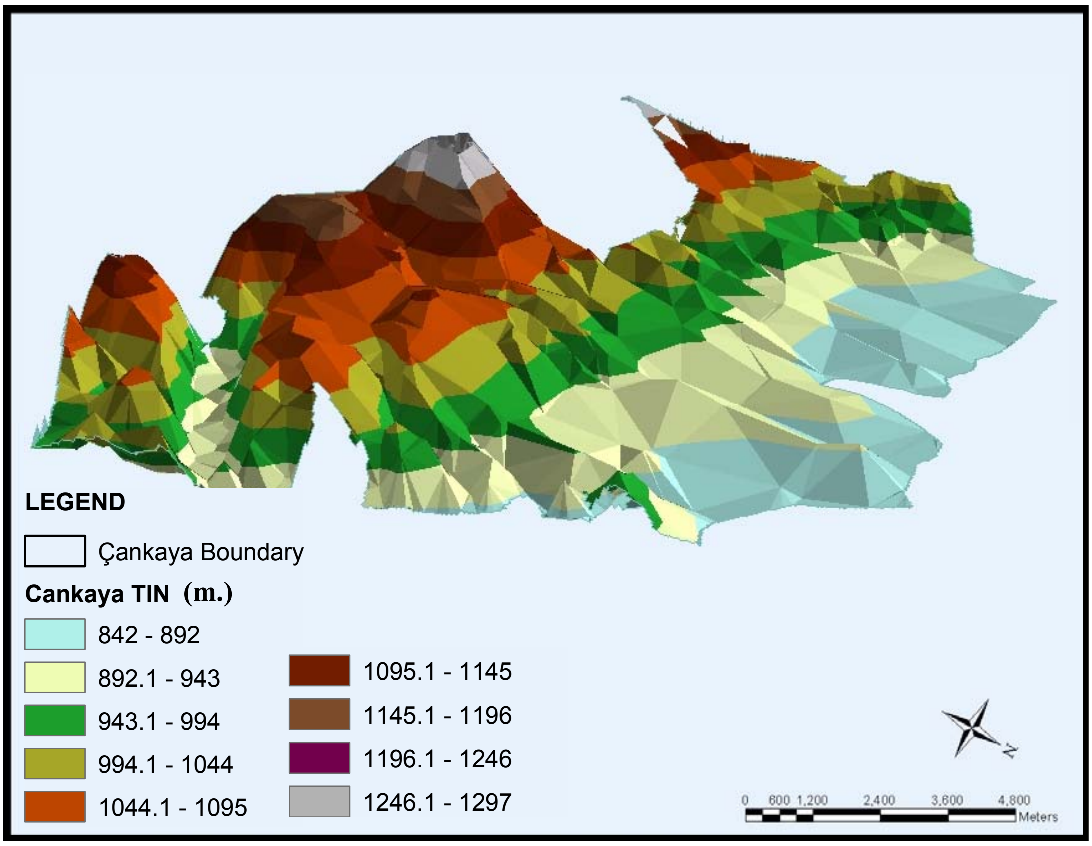


Figure 4.21 3-D representation of Çankaya district

### 4.4 Data Manipulation

#### 4.4.1 Data Conversion

After data collection period, ArcGIS 8.3 software has been chosen as the main software package; ArcView 3.2, MapInfo 7.5, MapBasic 5.5, and Visual Basic 6.0 have been used as complementary software packages for additional purposes.

ArcGIS 8.3 gives chance to use three different file formats (shapefile, coverage, geodatabase) for the non-imagery files. After learning the capabilities of each of these formats, and based on personal experiences, it has been understood that geodatabase format would be the most suitable one for this study because it supplies a uniform repository of geographic data, data entry and editing is more accurate, obtains more intuitive data objects, its features have a richer context, better maps can be made, features on a map display are dynamic, sets of features are continuous, and shapes of features are better defined (Chrisman, 1996). Therefore, all the digital data (shapefiles, mapinfo files) were needed to be converted to geodatabase format, and then all the tabular data were needed to be related to these geodatabases.

Shapefiles are converted to geodatabase format by the ‘shapefile to geodatabase’ conversion tool of ArcToolbox. To convert mapinfo files to geodatabase, firstly mapinfo files were converted to mif (mapinfo interchange file) format by the help of export tool of MapInfo software. Then, by the ‘MIF to Shapefile’ conversion tool of ArcToolbox, created mif files were converted to shapefiles. Then these shapefiles were converted to geodatabase format by the method described above.

Connecting tabular raw data in excel and access file formats to geodatabase files was not an automated process. All the tabular data had been prepared for the quarters, which were already in digital format, were needed to be connected to graphical quarter data. Firstly, new columns were added to the attribute table of ‘Çankaya Quarters’ geodatabase file by the names of the tabular data. There are only 104 rows in this table and all data were entered manually to the attribute table. In result, quarters have become more meaningful

### 4.4.2 Data Editing

Data editing operations have been made after conversion procedures that were done to study with a uniform data set of geodatabase files, which are more convenient for editing operations. First of all, this study needs a more detailed basic unit rather than quarters. The smallest spatial unit in the current statistical procedure is a quarter. Quarter is a rather large unit, and it does not give an opportunity for a detailed analysis about the different characteristics of inhabitants.

Therefore, 'block' has been selected as the smallest statistical unit after examining the United States (U.S.) and United Kingdom (U.K.) methodologies. A block is defined as an area bounded on all sides by visible and/or invisible features shown on a map prepared by the authorized organizations (US Census Bureau, 2000). A block will be the smallest geographic entity for which the State Institute of Statistics will tabulate decennial census data.

As it was already mentioned, Çankaya already has a block structure but this structure has many deficiencies. Therefore, a new block structure, which preserves the accurate blocks of the raw data, has been produced according to the rules described in Section 3.6.

Topology toolbar of ArcMap has been used to generate the new blocks and to revise the existing inaccurate blocks. Firstly, hundreds of road segments existing on unplanned areas were selected manually. Secondly, 'Construct Features' tool of the toolbar was activated. The cluster tolerance value was defined as 0.000032 automatically by the software. Finally, all the selected segments, that form a closed area when come together, were transformed to polygons, which would be accepted as the new blocks. Consequently, 3554 blocks in Çankaya district were created according to the rules defined. Figure 4.22 shows the disharmony of the old planned block structure with the current urban pattern while Figure 4.23 shows the harmony of new block structure in the same squatter settlement area.

Figure 4.22 Disharmony between the planned block structure and the current pattern

Figure 4.23 Proposed block structure for the same area shown in Figure 4.22

There were some digitizing errors derived while creating blocks, like dangles, switchbacks, slivers, knots, loops, under and overshoots etc. in the 'blocks' map. They had to be corrected in order to use these blocks as the basic units of the new census geography. Firstly, geodatabase file of blocks was converted to coverage format to be suitable for correction operations. Secondly, topology was built for the coverage by using the 'topology' tool of ArcToolbox software. Third step was to correct the errors in the coverage. 'Clean' command of the same tool was used for this purpose; location of coverage was given, dangle length and fuzzy tolerances (0.05) were entered, feature class was defined as poly and the location and name of the output coverage was given. After approving, all of the errors were eliminated.

Now all the data have been based on block level, decisions can be made according to this smallest unit. There are some 'small area estimation' techniques that could be used to estimate the values of the blocks according to their position in their quarter and according to their position in surrounding quarters. This study applies none of these methodologies because there is no information about general tendencies of the inhabitants in the quarters. For instance, it can not be said for any quarter that the quality of life is higher in certain part in the quarter or decreases through any direction in the quarter because our data are quarter based and no method will be meaningful to estimate the characteristics in each block.

It is still needed to assign the quarter based raw data, described in section 4.2, to their blocks. Therefore, the spatial join tool of ArcMap was used to join the geographical boundaries of the newly created block structure with the quarters. After this operation, all the raw data described till now were transferred to each of the blocks. Finally, the necessary block level data were ready for use to create the new census geography.

# CHAPTER 5

# METHODOLOGY

The standard census geography, proposed to be used in the 2010 Census, has already been defined in Chapter 3. Until now, boundaries of NUTS 1-2-3 and LAU 1 (Local Administrative Units) have been shown. This was the easy part of this thesis because the first three had already been defined and the last one has been created by a simple assumption. Now, it is time to create the other complementary parts of the proposed census geography, which are LAU 2, LAU 3 and LAU 4.

## 5.1 Finding homogeneous areas in the possible extent

Some general rules have already been defined about what conditions are needed to be fulfilled in order to define an area as a statistical census unit. Two of them are the most important ones. First, census units must have similar population characteristics, living conditions and housing characteristics to the possible extent. Second, the census units must be roughly equal in terms of population. That means, an optimum population criterion will be defined for each census unit and the total population of each unit should be as close as possible to these determined numbers.

Now the first step is to draw the boundaries of relatively homogeneous areas, which will be defined according to some criteria. Firstly quality of life degrees of the inhabitants will be determined, and then people in the same degree will be gathered in the same classes. Quality of life is described by Sactaqc (2005) as; “Those aspects of the economic, social and physical environment that make a community a desirable place in which to live.” Quality of life is a relative term and its definition may change according to the perspective of different people, who have different professions, cultures, stages of life etc. Therefore, deciding what makes a good life and what makes a better one is very difficult. Nevertheless, the United Nations Development Programme tries to determine the indicators of a high quality of life in their Human Development Report 2004. These human development indicators include 33 entries (see Appendix B) and sub-entries of them.

In this study, only ten of these indicators are used because some of them are not suitable for such a district scale study, and some of them need unavailable data. These ten indicators of quality of life are Literacy Ratio, University and Primary School Graduate Ratios (included in Education caption); Unemployment Ratio and Young Age Employment Ratio (included in Labor Force caption); Proximity to Police Stations, Hospitals, and Fire Stations (included in Proximity to Public Safety Services caption); Average Household Size, and Population Density.

Here, one of the most important indicators used in this study is ‘Labor Force’ including the ‘Unemployment Ratio’ and ‘Young Age Employment Ratio’ indicators. Income, based on employment, determines the power of a household to buy goods and services. So, unemployment directly affects housing, health, education and overall living qualities of household members. Also, Higher Young Age Employment Ratio is an indicator of low quality of life, too; because people in 12-19 years of age have to work in a job instead of studying in a school like others of the same age.

Another indicator is ‘Education’. Although this indicator is not directly effective on the level of quality of life, “…individual and community levels of education have a strong positive association with a range of economic and social benefits. An educated population adds to the vibrancy and creativity of their region” (NZCC, 2003). Furthermore, according to the results of ‘2002 Poverty Study’ report prepared by SIS and World Bank (2004), through the general perspective of Turkey, 41 % of illiterate people are poor while only 1,5 % of people graduated from universities, faculties etc.

live in the same economic conditions. The report concludes that “The higher the education level, the lower is the poverty risk in Turkey”.

The next indicator is ‘Average Household Size. Both ‘2002 Poverty Report’ of SIS and World Bank (2004) and ‘Poverty Assessment Report for Turkey – Volume 1’ of World Bank (2003) agree that “In Turkey, when the household size increases, the poverty rate increases, as well.” Of course, there are some exceptions of this finding such as single person households; and living this sort of lifestyle may be a choice enabled by income, not a necessity; or may be an obligation for divorced, widow, retired etc. people. SIS (2004) indicates that, “only 1 percent of the urban population lives in a single person household”, and it is an acceptable exception ratio for this study.

Although these reports are not directly related with quality of life concept, they give definite clues about the quality of life degrees of people in large households, illiterate people and people with high/low level of education. Therefore urban parts with higher average household sizes and lower education levels will be accepted as having lower quality of life than the others.

The next indicator is ‘Population Density’. High population density makes each person able to use a smaller share of the all types of natural and urban resource ‘pie’. Generally population density damages quality of life when highly increased. Nevertheless, in some cases, high density may become an important advantage to have more public services and social benefits, too. Higher population density will be accepted as an indicator of low quality of life; but because of its various effects on different population structures, the weight of population density will be smaller than the others to prevent making wrong evaluations.

Another indicator is ‘Proximity to Public Safety Services’. It is very important for inhabitants to feel safe. Therefore, proximity to fire stations, hospitals and police stations increase the safety feeling of people and make them more peaceful.

A Multi Attribute Decision Making (MADM) method will be used to combine all of these indicators and to find the quality of life degrees of living areas. The first step is to find the most qualified blocks and then produce a rank in a descending order beginning from the most qualified to the least. Firstly, the preferences have to be defined because they will be very effective when deciding on which and how characteristics would be more effective through the decision process. The MADM method will be performed from urban planning perspective and the conditions below will be accepted as the indicators of a high level of quality of life;

- Maximum literacy ratio
- Maximum university graduate ratio
- Minimum primary school graduate ratio
- Minimum unemployment ratio
- Minimum young age employment ratio
- Minimum population density
- Minimum average household size
- Maximum proximity to fire stations
- Maximum proximity to hospitals
- Maximum proximity to police stations

All of the necessary raw data including the values of these assumption criteria have already been defined in the attribute table of each block; so, the necessary data are ready to be used in a grading method. Before deciding the type of grading technique, an in-depth literature survey had been made, web sites of international census bureaus had been searched and many correspondences with their staff had been made about their applications and methodologies.

After the literature survey, it was certain that a multi-attribute decision making (MADM) method would be used for the application, because the aim in this step is to define a rank and arrange the living areas according to their grades. A multi-attribute decision rule is a process that allows for ranging alternatives. “It integrates the data and information on alternatives and decision maker’s preferences into an overall assessment of the alternatives and chooses the most preferred alternative to rank the alternatives in descending order of preference” (Malczewski, 1999, p.197). There are numerous MADM methods like Simple Additive Weighting (SAW), Value/Utility Function Approach, Analytic Hierarchy Process (AHP), Ideal Point Method, Fuzzy Aggregation Operations, etc. All of them have been examined and seen that each one has some special features and has the capacity to give different result. Two of them, SAW and AHP were tested in this study because of their simplicity and the types of the data needed for them. Simplicity is an important characteristic that makes these methods to be used widespread by further potential users.

At first, homogeneous areas were tried to be determined by the SAW method. In the application procedure, the main steps proposed by Malczewski (1999) were followed. As a result of the SAW process, a homogeneity map showing the boundaries of the living areas of people with similar quality of life degrees was drawn. Then, the same method was tried with different weights, and very different outputs were taken. It was understood that, the sensitivity of SAW method is very high, and it is incapable to meet the needs of such a complex project having many indicators. After the SAW method, Analytic Hierarchy Process (AHP) method was applied as the second alternative in the light of Malczewski’s (1999) opinions.

AHP procedure has three major principles; which are decomposition, comparative judgment, and synthesis of priorities. According to *decomposition principle,* the defined overall problem must be decomposed into its sub-parts in a hierarchical order and without missing the essential elements of the problem (see Figure 5.1). According to *comparative judgment* principle, all elements in the hierarchical structure must have pair-wise comparisons to each other, with respect to their parent in the next-higher level. According to the *synthesis principle,* the derived ratio-scale local priorities in the various levels of the hierarchy are combined to create a composite set of priorities for the elements at the lowest level of the hierarchy. Previous to the necessary calculations of AHP, a hierarchical decision-tree was formed. In decision-tree, "…the top level is the ultimate goal of the decision at hand, and the hierarchy descends from the general to the more specific until a level of attributes/sub-attributes is reached" (Malczewski, 1999).

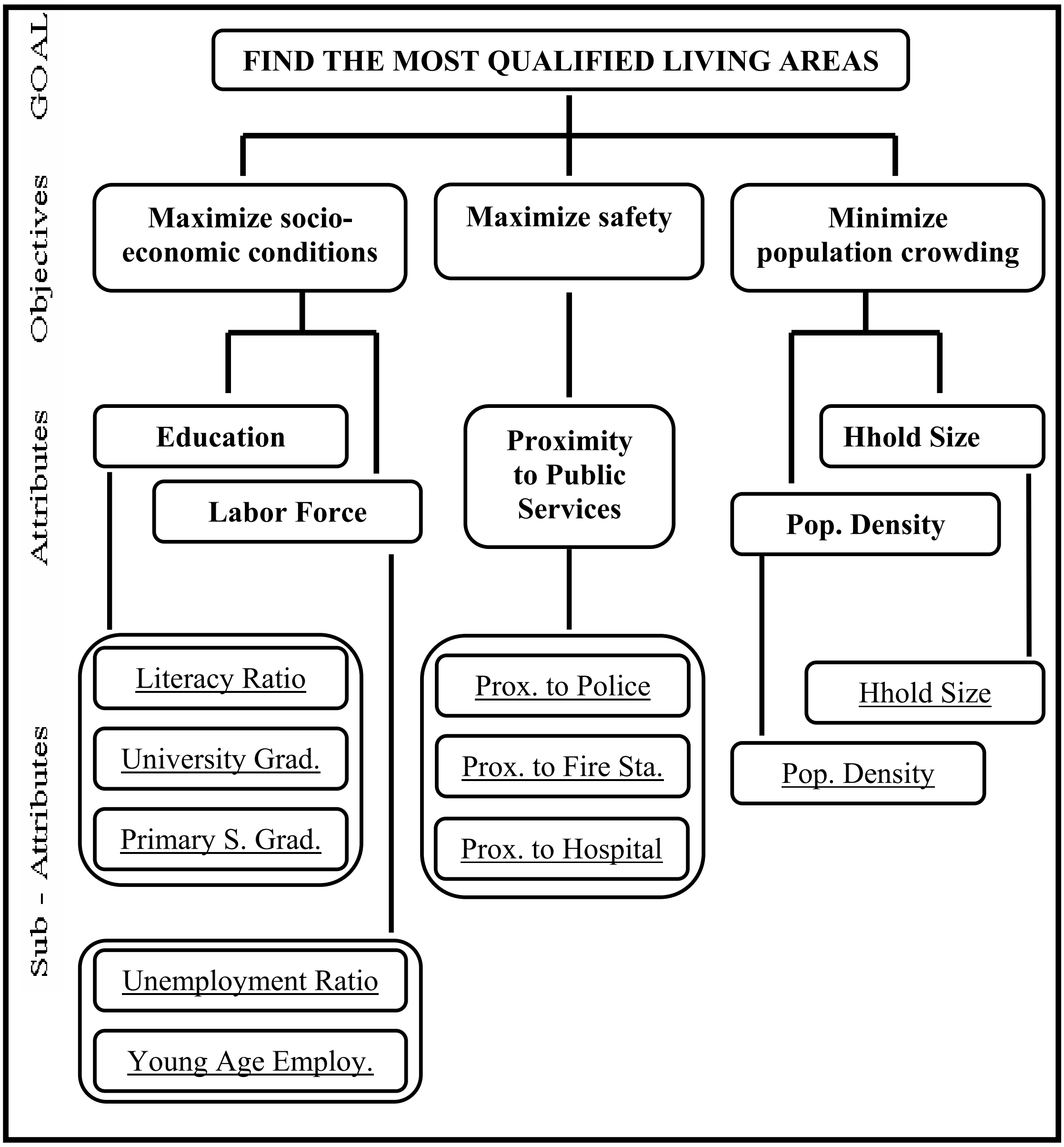

Figure 5.1 Analytic Hierarchy Process, Decision – Tree

The further steps of AHP procedure need some type of complex arithmetic calculations; and it is also necessary to make the analysis for several times. Therefore, a program is necessary to perform the necessary complex procedures. Finally, an application program including approximately 950 lines of code was written by using MapInfo 7.5, MapBasic 5.5 and Visual Basic 6.0 software packages. The program is an Integrated Mapping application, running on Visual Basic platform, and including MapBasic and MapInfo commands.

As it is seen in Figure 5.1, the ultimate goal is ‘to find the most qualified living areas’. 3 objectives, 5 attributes and 10 sub-attributes have been used to reach this overall goal. Each of the objectives has a different relative importance at each level of the hierarchy. Now, it is time to use the Visual Basic program to apply the Analytic Hierarchy Process and to find a rank decreasing from the most qualified to the least qualified living areas. The program starts with asking to select one of the two previously defined multi-attribute decision making methods, which are SAW and AHP. For further steps, the AHP method is selected from the first window of Graphical User Interface (GUI) shown in Figure 5.2.

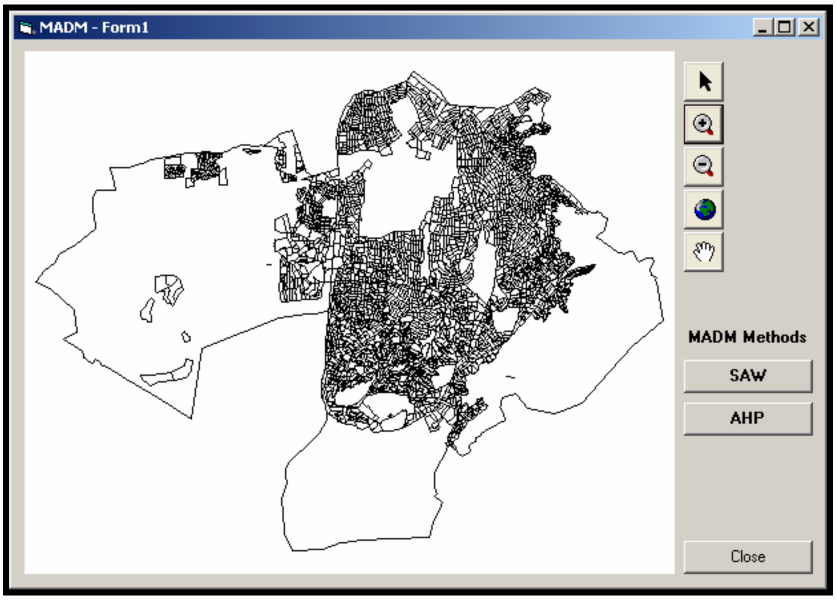


Figure 5.2 MADM – Selecting a method

The next step includes 'comparative judgment' (Pairwise Comparison Matrix), which is the single step that the user actively interferes to the process by stating his/her preferences and priorities. The user has to enter the importance of each attribute according to the others. As it is seen in Figure 5.3, there is a diagonal of textboxes from Northwest to Southeast, which include only the value of 1. This is a standard, because the importance of an attribute according to itself is always 1. The user is only responsible to fill the textboxes, which are at the Northeastern part of the diagonal. While filling them, user firstly reads the attributes at top row, and then defines for each attribute how many times more important than the attributes at the columns. The intensity of the importance value can be decided by looking to the "Scale for Pairwise Comparison Matrix", which gives the equivalent values of our comparisons in daily language (See Figure 5.3).

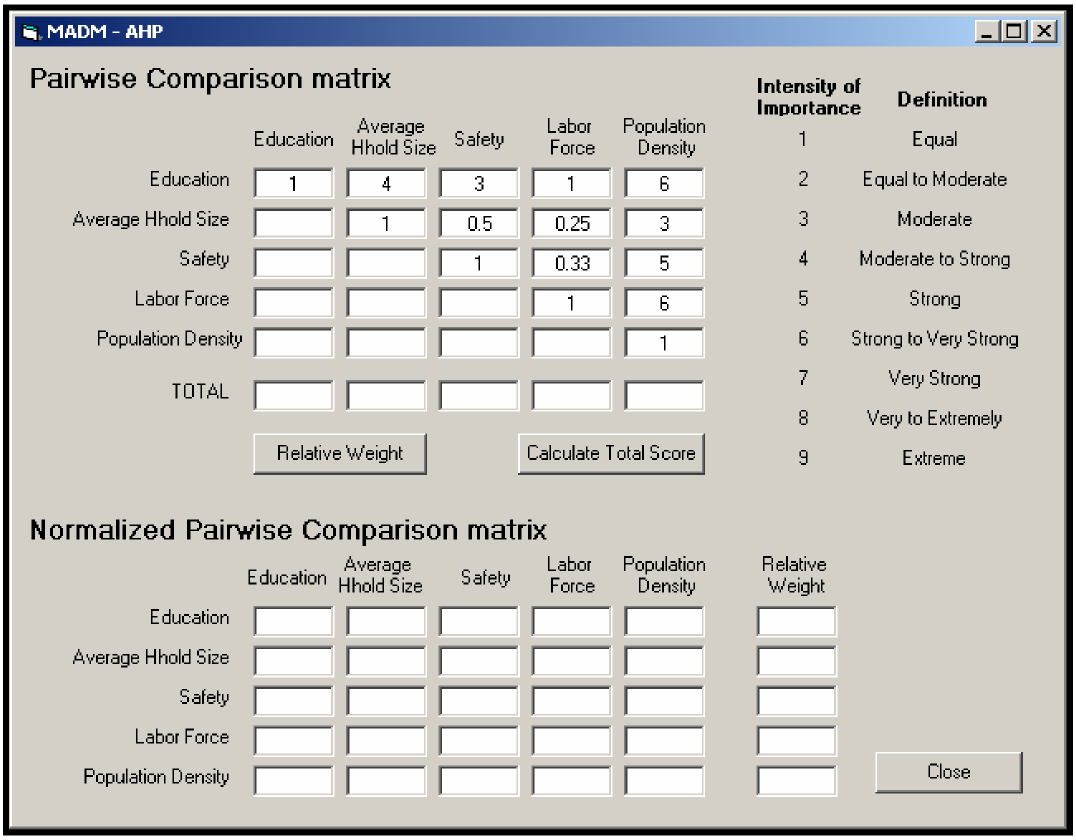


Figure 5.3 MADM – Comparative Judgment

All of the following steps in the program are done in control of the user. After giving the intensity of the importance values, user select the “Relative Weight” command and program calculates the values in all the empty textboxes in the table. 1 value will be divided by an entered value in the coordinate of “B column x A row” of the matrix and the result will be entered by the program to the “A column x B row”. Then, for future needs, total values of each column are calculated and written to their places in the row named TOTAL.

The next step is the normalization of the results of pairwise comparison matrix. The same numbers of textboxes are again in the same coordinates. Normalized value of each of the values seen at the upper matrix is written to these textboxes. Normalization process includes the division of each cell value in the upper matrix by the TOTAL of its column values. Then the results are automatically entered to the cells in the matrix. Almost all of the results become smaller than 1 and hard to perceive. Therefore, each of them is multiplied by 1000 to make them more understandable. This operation will have no effect on the results.

The final step is to calculate the relative weights of each attribute. To do that, the average value of each row in the “Normalized Pairwise Comparison Matrix” is calculated and written to the textboxes of the “Relative Weight” column at the end. The result of these processes according to the previously defined preferences and priorities is shown in Figure 5.4.

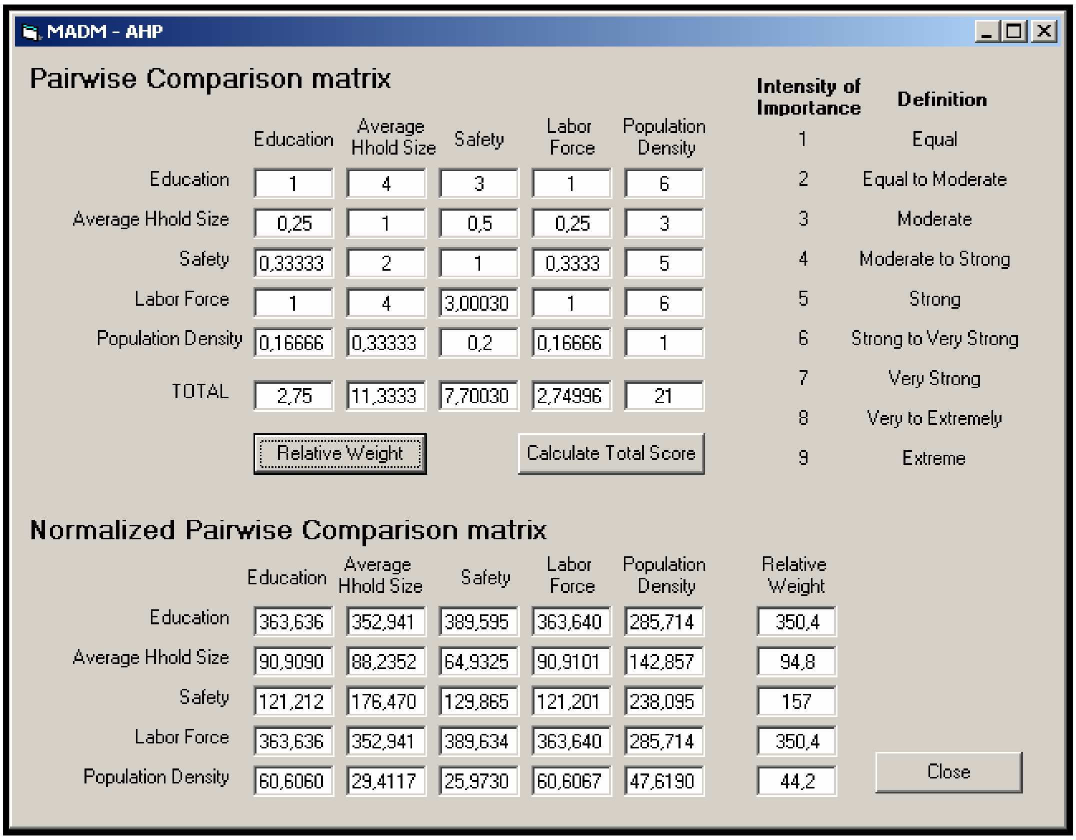


Figure 5.4 MADM – Calculation of relative weight

After calculating relative weights, the user selects the "Calculate Total Score" command and program automatically begins to perform the necessary operations to find out the score of each class of each criterion. Here the term 'class' means an interval in a criterion defined according to the natural distribution of the values of that criterion. The AHP process contains 10 criteria and each of these criteria is divided into 6 classes (first class is assigned to not applicable areas and not included in the scoring procedure), each of which was scored according to the assumptions already stated. Then the scores were standardized by the score range procedure, using linear scale transformation formula;

$$X'_{ab} = (X_{ab} - X_{min.b}) / (X_{max.b} - X_{min.b}) \qquad (5.1)$$

Where $X'_{ab}$ is the standardized score for the *a*th object and the *b*th attribute, $X_{ab}$ is the raw score, $X_{min.b}$ is the minimum score for the *b*th attrribute, $X_{max.b}$ is the maximum score for the *b*th attrribute, $X_{max.b} - X_{min.b}$ is the range of a given criterion. “The standardized scores are calculated by dividing the difference between a given raw score and the minimum score for the criterion by the score range for the benefit criterion, and the standardized score is obtained by dividing the difference between the maximum score and a given raw score by the score range” (Malczewski, 1999). The values of standardized scores can range between 0 and 1 as shown in Table 5.1.

Table 5.1 Scores and standardized scores of each criterion used in AHP procedure.

| **Literacy Ratio** (%) | **Score** | **Std.Score** | **Avg. Household Size** | **Score** | **Std.Score** |
|---|---|---|---|---|---|
| Not Applicable | 0 | 0 | Not Applicable | 0 | 0 |
| 89.3 – 92.3 | 1 | 0.2 | 2.05 – 2.29 | 5 | 1 |
| 92.4 – 94.6 | 2 | 0.4 | 2.30 - 2.55 | 4 | 0.8 |
| 94.7 – 96.2 | 3 | 0.6 | 2.56 – 2.76 | 3 | 0.6 |
| 96.3 – 97.6 | 4 | 0.8 | 2.77 - 3.05 | 2 | 0.4 |
| 97.7 – 99.3 | 5 | 1 | 3.06 - 4.63 | 1 | 0.2 |
| **Unemployment Ratio** | **Score** | **Std.Score** | **Population Density** | **Score** | **Std.Score** |
| Not Applicable | 0 | 0 | Not Applicable | 0 | 0 |
| 3.8 – 8.8 | 5 | 1 | 49 - 68 | 5 | 1 |
| 8.9 – 10.5 | 4 | 0.8 | 69 - 101 | 4 | 0.8 |
| 10.6 – 11.5 | 3 | 0.6 | 102 - 148 | 3 | 0.6 |
| 11.6 – 12.6 | 2 | 0.4 | 149 - 193 | 2 | 0.4 |
| 12.7 – 17.5 | 1 | 0.2 | 194 - 468 | 1 | 0.2 |
| **Young Age E. Ratio** | **Score** | **Std.Score** | **Proximity to Hospitals** | **Score** | **Std.Score** |
| Not Applicable | 0 | 0 | 1 min. | 5 | 1 |
| 1.8 – 3.3 | 5 | 1 | 3 min. | 4 | 0.8 |
| 3.4 – 5.7 | 4 | 0.8 | 5 min. | 3 | 0.6 |
| 5.8 – 6.9 | 3 | 0.6 | 8. min | 2 | 0.4 |
| 7.0 – 11.1 | 2 | 0.4 | 10 min. | 1 | 0.2 |
| 11.2 – 20.2 | 1 | 0.2 | 10 min.+ | 0 | 0 |
| **University Gra. Ratio** | **Score** | **Std.Score** | **Proximity to Police Sta.** | **Score** | **Std.Score** |
| Not Rated | 0 | 0 | 1 min. | 5 | 1 |
| 10.5 – 14.1 | 1 | 0.2 | 3 min. | 4 | 0.8 |
| 14.2 – 23.5 | 2 | 0.4 | 5 min. | 3 | 0.6 |
| 23.6 – 29.0 | 3 | 0.6 | 8. min | 2 | 0.4 |
| 29.1 – 39.1 | 4 | 0.8 | 10 min. | 1 | 0.2 |
| 39.2 – 48.0 | 5 | 1 | 10 min.+ | 0 | 0 |

Table 5.1 Scores and standardized scores of each criterion used in AHP procedure (cntd)

| **Primary School Graduate Ratio** | **Score** | **Std.Score** |
|---|---|---|
| Not Rated | 0 | 0 |
| 11.2 – 18.8 | 5 | 1 |
| 18.9 – 29.4 | 4 | 0.8 |
| 29.5 – 35.4 | 3 | 0.6 |
| 35.5 - 47 | 2 | 0.4 |
| 47.1 – 76.7 | 1 | 0.2 |

| **Proximity to Fire Sta.** | **Score** | **Std.Score** |
|---|---|---|
| 1 min. | 5 | 1 |
| 3 min. | 4 | 0.8 |
| 5 min. | 3 | 0.6 |
| 8. min | 2 | 0.4 |
| 10 min. | 1 | 0.2 |
| 10 min. + | 0 | 0 |

Std. Score means Standardized Score of the class.

Next, **weighted scores** are calculated by multiplying the previously calculated "Relative Weights" and "Standardized Scores" of each criterion (Figure 5.5).

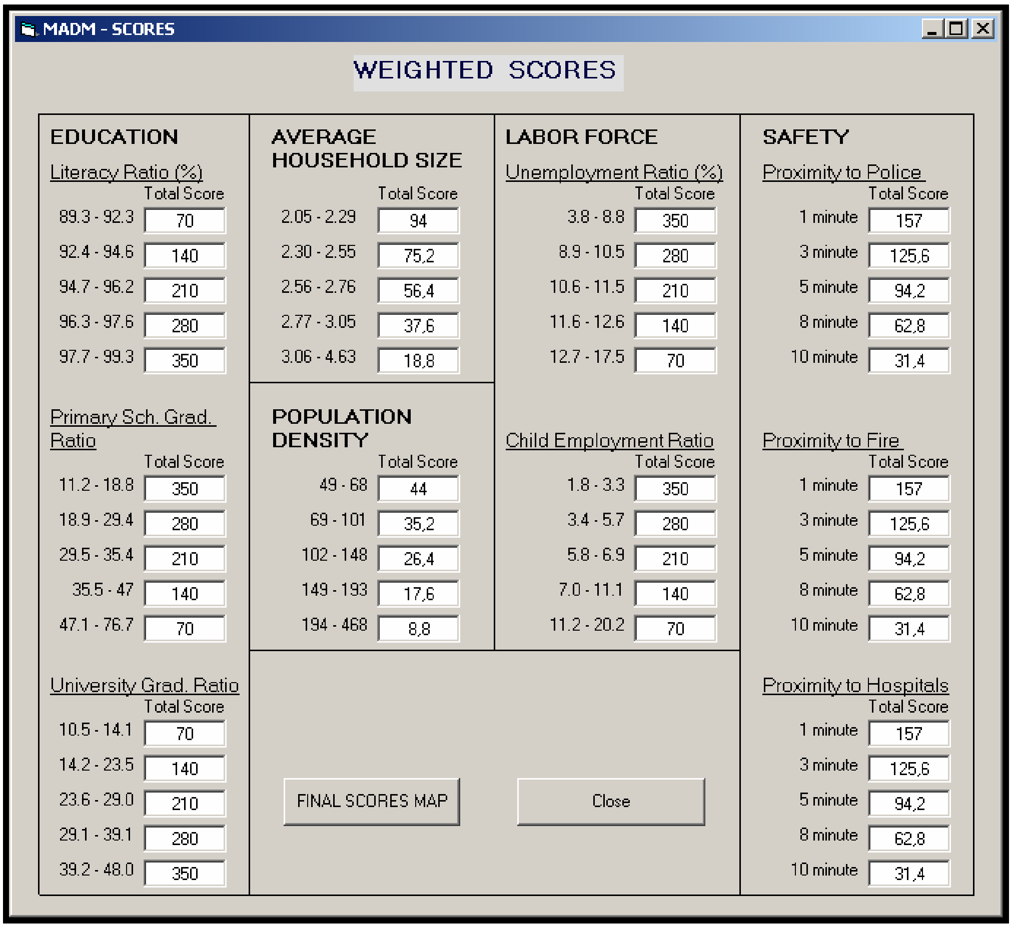


Figure 5.5 Weighted Scores of the indicator ranges

The next step is to assign these scores to the smallest statistical units, blocks. Firstly, the table that represents all of the blocks is updated by adding 11 new fields to its attribute table. 10 of these fields are for the scores of each criterion and the last one is to be filled with the summation of the scores of indicators for the related block. Therefore, the "blocks" table is opened in the background of the program, and then score of each block is calculated to each criterion according to the state of belonging of the block to the classes already defined. Then, all of the score values in the 10 field of each block are summed up, and the overall scores are obtained in a single field, named "Total Scores". Now, all of the blocks have their relative scores ranging between the most qualified block/blocks and the least qualified block/blocks. The final step is to create a thematic map from all of the individual values by assigning a different color for each of them. The user selects firstly the "Final Scores Map" command, shown in Figure 5.5, and a new form is opened. Here 'Thematic Map' button is clicked and the program creates a thematic map of areas as homogeneous as possible according to their quality of life degrees (see Figure 5.6).

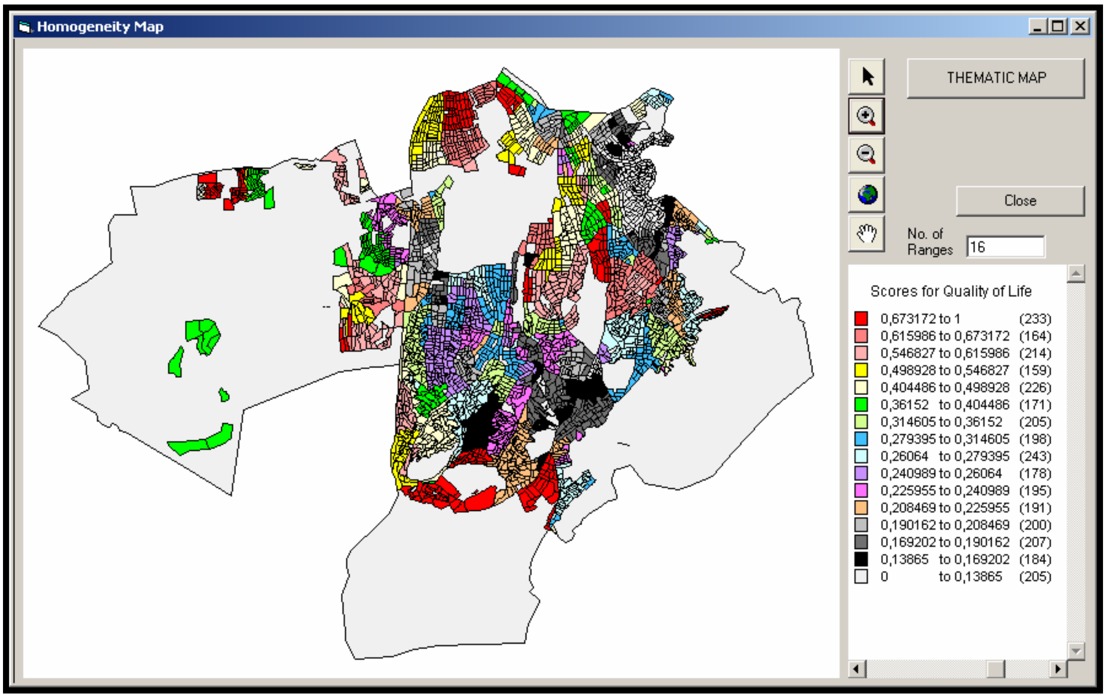


Figure 5.6 Homogeneity Map according to quality of life degrees

### 5.2 Sensitivity Analysis

"Sensitivity analysis examines the sensitivity of the results to changes in the priorities of the criteria. This is a particularly important aspect of an AHP problem analysis, since results are based on subjective expert assessments" (Abel, 1996). In AHP procedure, it may be applied in any step. The importance of one or more indicators is changed, then the changes in the outcomes are examined, and finally a comparison is made. In this way, the most effective indicators are determined, too. "Where the project is shown to be sensitive to the value of a variable that is uncertain, that is, where relatively small and likely changes in a variable affect the overall project result, mitigating actions at the model should be considered" (ADB, 2005).

The main purposes of sensitivity analysis are to assess the stability of the optimal solution under possible changes in parameters, the impact of the lack of controllability of certain parameters, and the need for precise estimation of parameter values. Sensitivity analysis can show the robustness of the overall priority rating, and to what extent the overall priorities are sensitive to changes in the importance of criteria. Such analyses enlarge both the understanding of and the confidence in the outcome of the AHP analysis. Sensitivity analyses also show the consequences of changes in, for example, the importance of indicators. Confidence in the outcome of the analysis will increase if small changes in the relative importance of factors do not have much impact on the overall priority rating. When the AHP analysis has been completed it is rather easy to determine the consequences of changes in the judgments on the overall priorities. A software package is required to perform the necessary complex operations.

In this study, the sensitivity analysis was first carried out by the perspective of the author, an Urban Planner, with the aid of the EXPERT CHOICE software; but different decision makers may have different thoughts and may give various importance values to the indicators. Therefore, by maximally changing the relative importance of each criterion one at a time, relative changes in the weights of the other indicators and in the global weights were observed. In this way, the maximum and minimum deviations of the model were determined, and the confidentiality of the model was tested. Through the sensitivity analysis, it would be very time consuming to try every possible combination. Instead, a one-way sensitivity analysis was performed.

In one-way sensitivity analysis objectives' weights, single attribute value functions, or attribute ratings for decision alternatives are varied, one at time, to see how sensitive the model is to those changes. The total values of decision alternatives are drawn as a function of the variable under consideration (HUT, 2002). Only ten different combinations, which include maximum (+9) and minimum (-9) values of five criteria, were performed. For example, according to an anonymous user, if education is 9 times more important then the other four criteria, it is a limit. In the same way, if safety is 9 times less important than the others, it is a limit, too. By using Expert Choice, the ten different applications were performed. The results of these operations, the reflections of the changes in the weights to the weights of the other indicators are shown in Table 5.2, and the global weights for each alternative are shown in Table 5.3.

Table 5.2 Sensitivity Analysis results showing the changes in the weights

| **ALTERNATIVES** | **Education** | **Avg.Hhold** | **Safety** | **Labor Force** | **Pop. Density** |
|---|---|---|---|---|---|
| **Original** | **0.354** | **0.094** | **0.156** | **0.354** | **0.044** |
| **Max. Education** | 0.666 | 0.055 | 0.086 | 0.163 | 0.03 |
| **Min. Education** | 0.024 | 0.161 | 0.252 | 0.475 | 0.088 |
| **Max.Avg.Hhold Size** | 0.122 | 0.664 | 0.065 | 0.122 | 0.028 |
| **Min. Avg.Hhold Size** | 0.354 | 0.024 | 0.187 | 0.354 | 0.081 |
| **Max. Safety** | 0.128 | 0.05 | 0.665 | 0.128 | 0.029 |
| **Min. Safety** | 0.372 | 0.146 | 0.024 | 0.372 | 0.086 |
| **Max. Labor Force** | 0.163 | 0.055 | 0.086 | 0.666 | 0.3 |
| **Min. Labor Force** | 0.475 | 0.161 | 0.252 | 0.024 | 0.088 |
| **Max. Pop.Density** | 0.116 | 0.037 | 0.053 | 0.116 | 0.677 |
| **Min. Pop.Density** | 0.351 | 0.112 | 0.16 | 0.351 | 0.025 |

Table 5.3 Sensitivity Analysis results showing the changes in the global weights

| ALTERNATIVES | Inconsistency | Min. | Original | Max. | Difference from Original | Difference ( % ) |
|---|---|---|---|---|---|---|
| Max.(+9) & Min.(-9) **Education** | 0.12 | 0.407 | 0.5 | 0.593 | 0.093 | 18.6% |
| Max.(+9) & Min.(-9) **Avg.Hhold Size** | 0.13 | 0.478 | 0.5 | 0.522 | 0.022 | 4.4% |
| Max.(+9) & Min.(-9) **Safety** | 0.12 | 0.463 | 0.5 | 0.537 | 0.037 | 7.4% |
| Max.(+9) & Min.(-9) **Labor Force** | 0.12 | 0.407 | 0.5 | 0.593 | 0.093 | 18.6% |
| Max.(+9) & Min.(-9) **Pop.Density** | 0.07 | 0.49 | 0.5 | 0.51 | 0.01 | 2.0% |

The sensitivity analysis results show that overall conclusions obtained are not fragile to the assumptions made regarding maximum elasticities, and the variations are in an acceptable range. All the global weight variations have been in the range of 2.0% and 18.6 %; and these are not large variations considering the scope of the study and the given maximum and minimum importance values. It was also seen that, weight changes in Education and Labor Force have a significant effect on the quality of life levels of the people.

In addition, inconsistency in Table 5.3 is a measure used in Expert Choice when making paired comparisons. An inconsistency ratio is calculated for each set of judgments. Inconsistency follows the transitive property, for example, if you were to say that A > B, and B > C, then say that C > A, you would have been inconsistent. Since a set of perfectly consistent judgments produces a consistency index of 0, the consistency ratio will also be zero. A consistency ratio of 1 indicates consistency akin to that which would be achieved if judgments were made at random rather than intelligently. This ratio is called the inconsistency ratio in Expert Choice, since the larger the value, the more inconsistent the judgments.

To measure the inconsistency of all the judgments made in the decision hierarchy, we take the inconsistency value of each set of comparisons and multiply it by the priority of the element with respect to which the comparisons are made, and add for all the elements. This gives a single overall weighted number. To decide how acceptable this number is, we form a ratio with a similar number obtained by multiplying the corresponding random inconsistency value for an equal number of comparisons by the priority of the elements, and again add over each attribute. According to Abel (1996), the resulting ratio should be 0.10 or less. Inconsistency ratios in Table 5.3 are the results of extreme weightings, so the 0.12 and 0.13 ratios are acceptable deviations for this study.

The sensitivity of the model was also tested on the final homogeneity maps, which are the results of minimum and maximum importance changes in the indicators. 10 different maps were created and changes were observed. None of them caused considerable changes on the results. Here only the result map of Education, which shows greater changes than the others, will be examined.

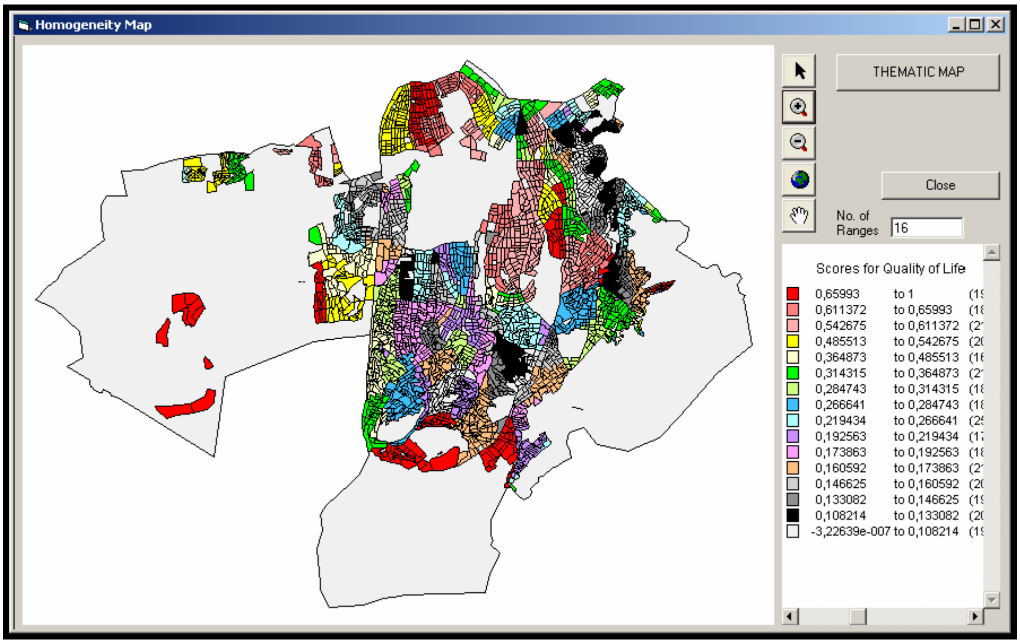


Figure 5.7 Homogeneity Map after giving maximum importance to Education

When Figure 5.6 (the map which will be used in further steps) and Figure 5.7 are compared to each other, it is seen that mapping results are parallel to the observed changes in the global weights, shown in Table 5.3. There are small changes in the differentiations of the quality of life levels across the case study area. According to the comparison results, only some of the blocks changed their rankings in the quality of life grading. Here the important point is that, the top levels and the bottom levels show almost no changes, and save their ranks. The changes generally occur in the medium parts of the grading and the blocks located here move either to top or bottom for 2 or 3 rows. This situation shows that, it is impossible for a squatter settlement area to move to the much upper levels of the grading by the change of importance of any indicator; because quality of life indicators are consistent to each other and none of them is completely determinant on the rank of a block by itself. That is, the maximum changes in the importance of indicators do not dramatically affect the overall project result. As a result, it can be said that the AHP model used in this study is robust against the probable subjectivities and has a highly confidential and consistent structure.

### 5.3 Creating LAU 2, LAU 3 and LAU 4 regions

The homogeneity map generated by integrated mapping program was in MapInfo format. This file was needed to be converted to shapefile format. Briefly, mapinfo files were exported to mif (mapinfo interchange file) format. Then, in ArcToolbox, after selecting ‘Import to Shapefile’ toolbox ‘MIF to Shapefile’ conversion tool was selected. Finally, the shapefiles were converted to geodatabase format to be used in districting operations.

The boundaries of relatively homogeneous areas were determined and the necessary conversion operations were completed. The next step was to transform the homogeneous areas into a meaningful and hierarchical structure. According to the proposed standard census geography hierarchy described in section 2.3, it was needed to create LAU 2, LAU 3 and LAU 4 regions in addition to the existing and assumed NUTS regions. After the previously defined homogeneity criterion, “Population” was the second criterion in forming the new LAU regions. As it was mentioned, the census geography units in the same level of hierarchy have to have approximately same number of inhabitants in order to be comparable to each other. The LAU regulation will lay down the optimum, minimum and maximum population thresholds shown in Table 5.4;

Table 5.4 Proposed population sizes of new LAU regions

| CLASS | Optimum | Minimum | Maximum |
|---|---|---|---|
| LAU 2 | 15,000 | 10,000 | 20,000 |
| LAU 3 | 5,000 | 3,000 | 7,500 |
| LAU 4 | 1,500 | 800 | 2,500 |

The next step was to divide the homogeneous areas according to the given population sizes. The software, which was used to create the boundaries of census geography units, is the “Districting” module of ArcGIS 8.3. Classification procedure with Districting Add-on included some specific steps. Firstly an empty document was opened in ArcMap. The Districting add-on was activated and added among the other toolbars in use. Using this toolbar, Districting Wizard was opened and the necessary information about source data was entered to “Create Plan” user interface.

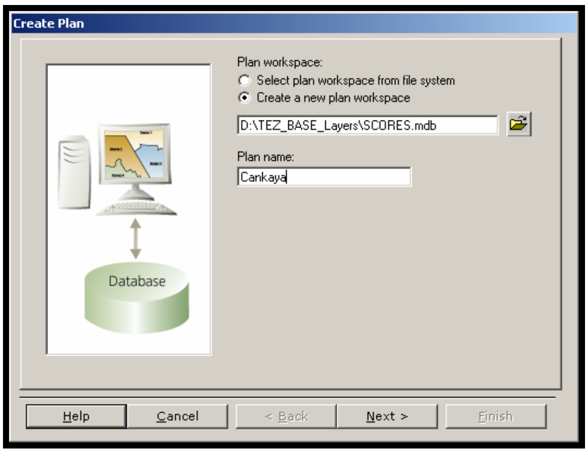


Figure 5.8 Creating a new plan workspace

Then, the "Scores" table having the scores of each block according to their quality of life degrees was entered as the location of source geography (see Figure 5.9). In addition, "source geography has attribute data" checkbox was checked. After checking this box, source geography and its attributes were going to be saved in the same file. In the third window, characteristics of the source geography were described. A plan workspace was selected to add the source geography; geographic data source type was defined; source geography data source was entered; source geography key field name was given; and the attribute data fields were selected as "Population", because in this case the classification will be done according to "Population" field (see Figure 5.10).

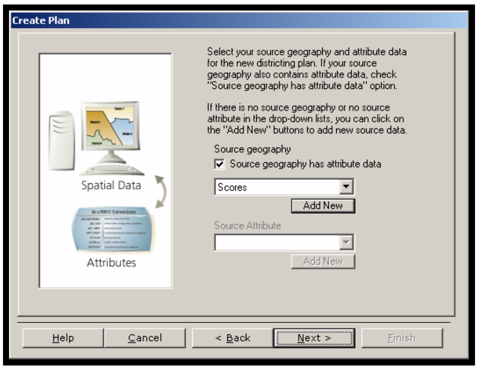


Figure 5.9 Defining the source geography

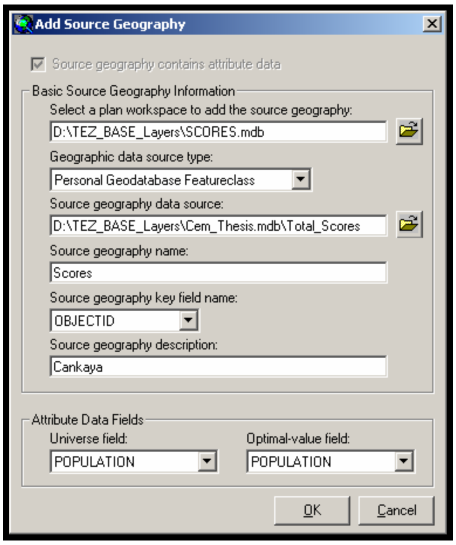


Figure 5.10 Adding the characteristics of source geography

Then, district caption was selected as 'District', and the preferred number of districts was selected from the given values ranging between 1 and 200 (see Figure 5.11). It is not allowed to add extra districts after beginning to create districts. Therefore, it was better to choose the highest number, 200. After the districting process completed, the excess districts were removed. Then, districting wizard gave the optimal target population value by dividing the total population to selected number of districts (see Figure 5.12).

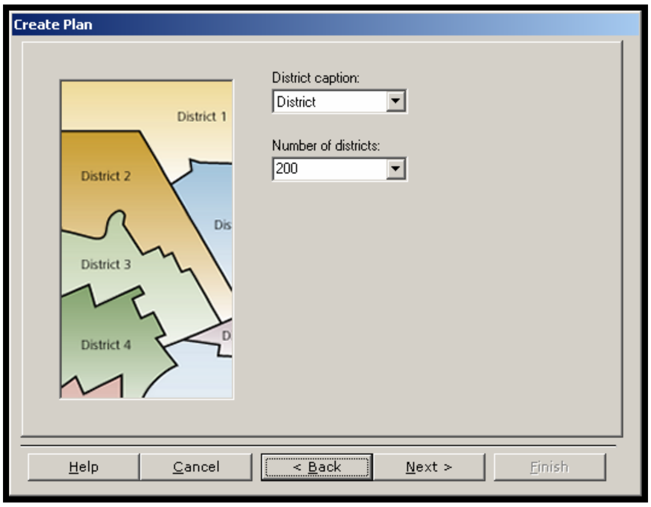


Figure 5.11 Deciding to numbers of districts

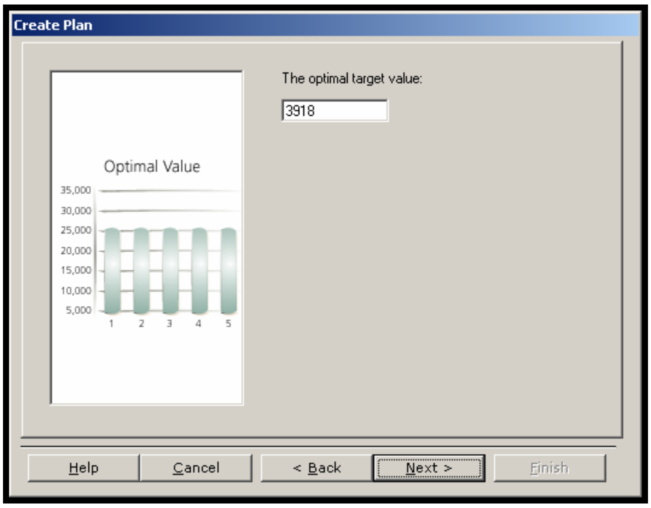


Figure 5.12 Calculated optimal target (population) value

The proposed optimal target population value was accepted. Then the program offered a ‘check list’ to check the fields to calculate sum and percentage values for each district (see Figure 5.13). Some necessary fields were checked here to see the statistical properties of generated districts after the process. Finally, after selecting “Finish” command, the program creates a document which includes unassigned districts, a statistics window mainly used to see the populations of created districts, a chart window showing the condition of population of each district according to the others, and a Districting toolbar to perform the necessary clustering operations (see Figure 5.14).

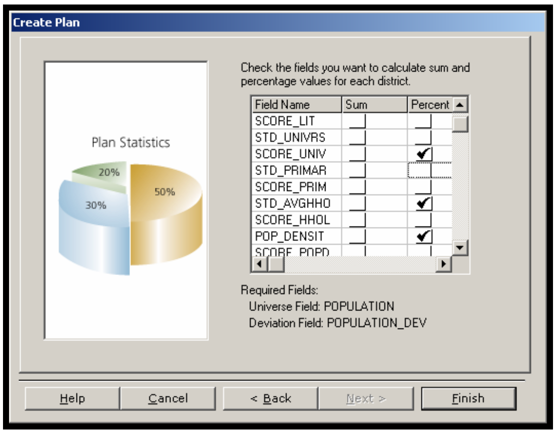


Figure 5.13 Choosing calculations from the check list

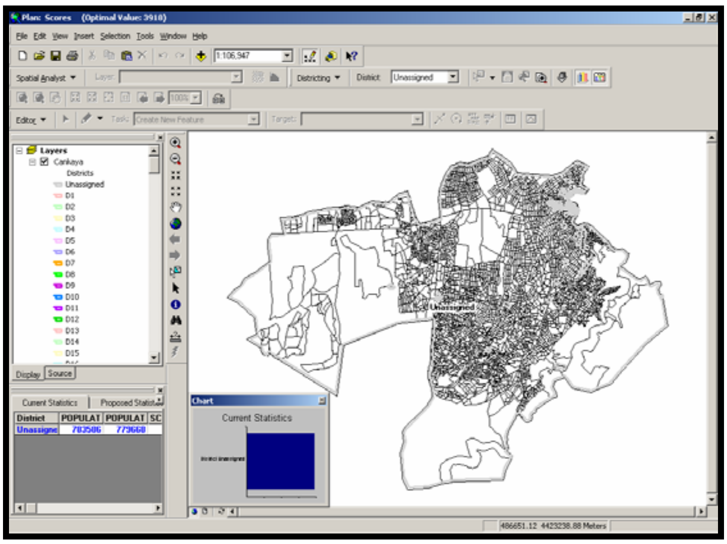

Figure 5.14 Districting working platform

During the districting process; blocks are selected manually. The “current statistics” table of the extension shows total population of the selected districts and how many inhabitants more are needed to arrive the optimum population size. In addition, the chart shows the population situation of that district according to previously created ones. If the user decides that the population is okay, then by the help of the toolbar, he/she assigns the blocks into a unique district.

Of course, some base maps should be used to decide which blocks should be selected to make accurate clustering operations. The most important of them is ‘Homogeneity Map’ created from the ‘Scores’ file produced as a result of MADM process. This map shows the boundaries of areas that are as homogeneous as possible. The blocks included in the equally scored and colored areas, are tried to be clustered into the same district during the process.

The second base layer is ‘Buildings’, which has been obtained as raw data and shows the locations of buildings. It is obvious that the homogeneity map can not be 100% accurate and some other data should be used to increase its accuracy. Buildings layer becomes important in this sense, because it will be used as a complementary map in the decision making process. The building patterns of the blocks and the block groups will be interpreted by visual examination. This method will dramatically increase the accuracy degree because of some known facts about the squatter settlements. They show irregular building pattern, and almost all of their inhabitants have low quality of life levels. After the integration of homogeneity and building pattern maps, the results will be more accurate when deciding on which blocks should be clustered together.

The third base layer used in districting process is ‘Transportation Hierarchy’. It is assumed especially for boulevards and avenues that the people living along the same roads have similar living conditions. Then, the blocks along the same road are tried to be clustered together if they are also in the same color in the homogeneity map. Slope map was also used as a base map sometimes to see the important separations caused by slope; but it was not as effective as the other base maps in the decision making process.

It was also considered to make some additions to the current program written in Visual Basic to make all the districting process automatically; but it is impossible for a computer to interpret the building patterns as accurate as human eye. Therefore, the MADM program has only been used to create the homogeneity map according to the user's preferences and priorities.

After all the preparations completed, the new LAU regions were started to be created. The first step was to create LAU 4 regions, each of which is planned to include optimum 1500 inhabitants. Then the upper LAU 3 and LAU 2 regions would be determined according to the boundaries of LAU 4 regions. Each LAU 3 region will include three LAU 4 regions, and each LAU 2 will include three LAU 3 regions.

LAU 4 regions have been created by the help of the base maps and according to the population criterion. If Figure 5.12 is carefully examined – "Calculated optimal target (population) value", it is seen that the optimum value was calculated as 3918, which is the result of the division of population of Çankaya by 200 districts. The proposed ideal number of inhabitants for a LAU 4 region is 1500 in this study, and total number of LAU 4 regions will be 461; but as a drawback of the program, the maximum district number to be created in a single plan is limited with 200. Therefore, the calculated value had to be accepted and LAU 4 districting process was executed on three different plans, which have separately 200, 200 and 60 regions. Later, these separate plans were opened in the same workspace and merged to each other by the help of "GeoProcessing Wizard" tool of ArcMap. The results are shown in Figure 5.15 below.

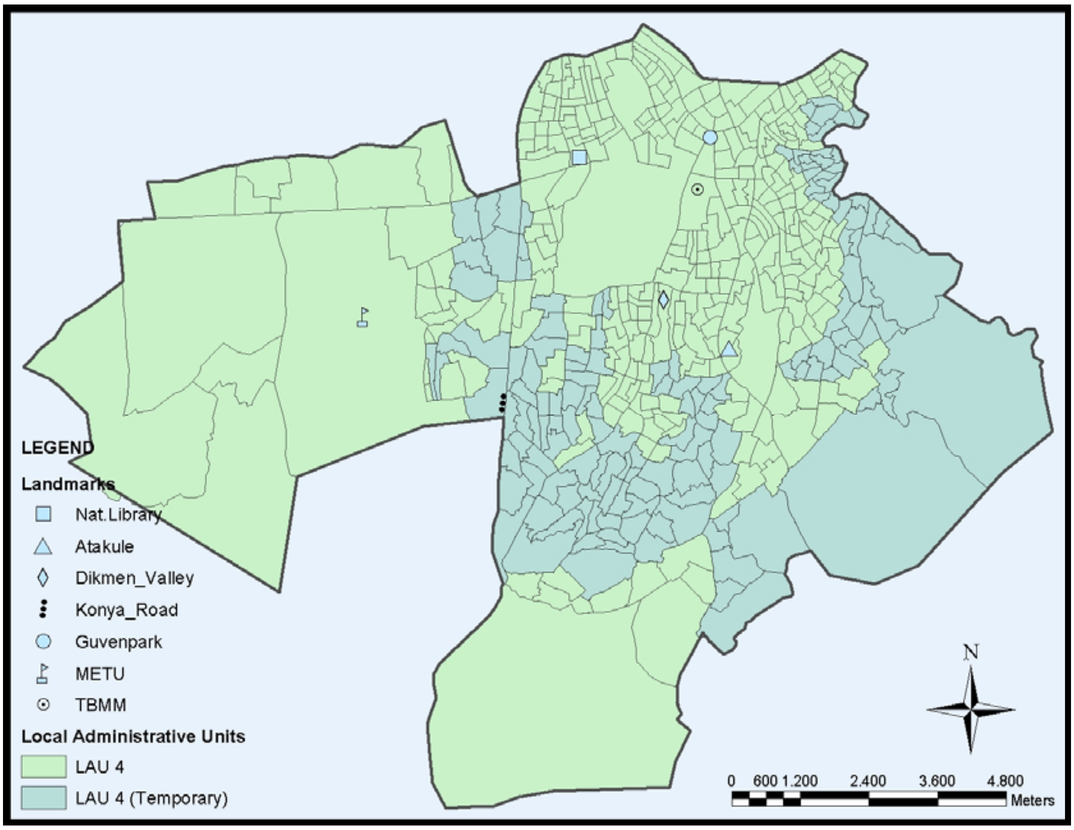


Figure 5.15 Permanent (313 units) and Temporary (148 units) LAU 4 regions

As it can be seen in the Figure 5.15, there are two types of LAU 4 regions in Çankaya, permanent and temporary LAU 4 regions. One of the basic rules while creating these districts was "to create census geographies which will not change over time", but Çankaya district includes some areas which are amenable to change radically in a short period of time. The temporary LAU 4 regions are generally formed by squatter settlements, which are likely to undergo a rebuilding process. Their both structural and population characteristics may change because of possible renewal plans, so their statistical boundaries may change, too. Because of these reasons, these areas were named as "Temporary LAU 4" regions. When they complete their rebuilding process, their boundaries will be redrawn and they will be located among the other permanent LAU 4 regions.

The next step was to create LAU 3 areas which include optimum 5000 inhabitants and generated by merging the LAU 4 regions. There are 159 LAU 3 regions proposed and each of them generally includes three units of LAU 4 regions (see Figure 5.16).

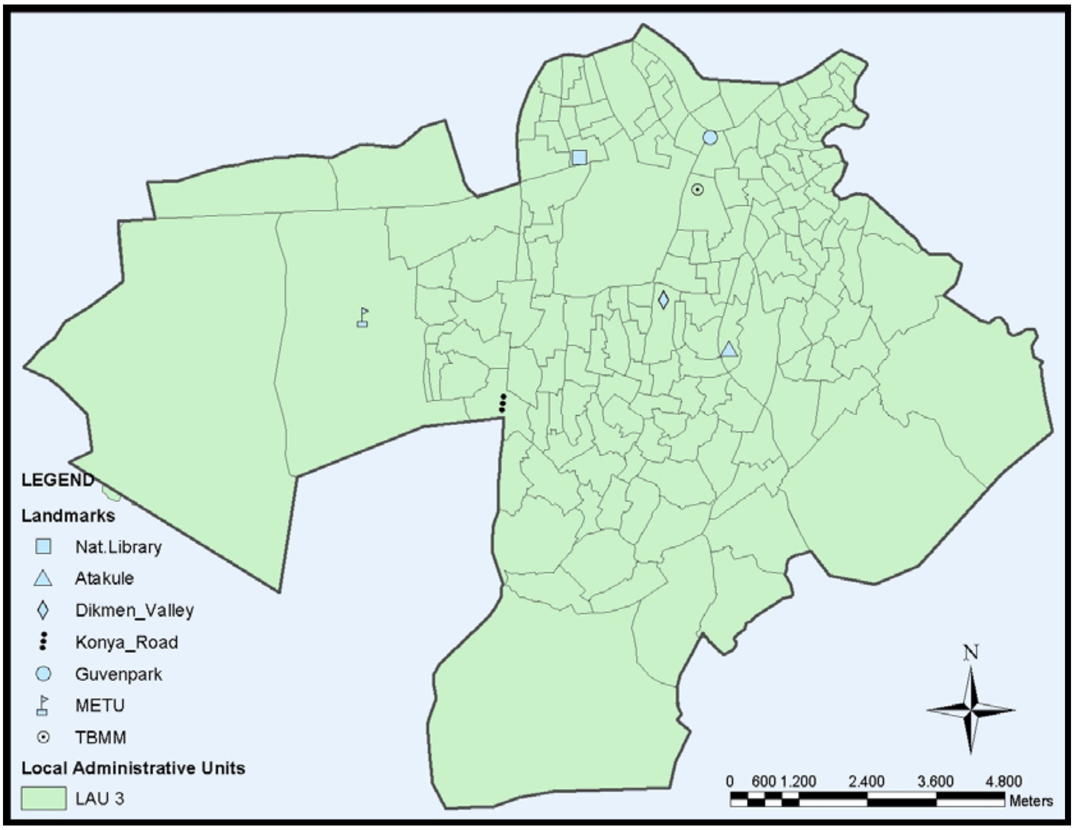


Figure 5.16 Boundaries of LAU 3 regions (159 units)

The final step to complete the standard census geography was to create LAU 2 areas, which are the upper level of regions of LAU 3 regions. They optimally have 15,000 inhabitants and three LAU 3 regions. There are 54 LAU 2 regions, which are considered to substitute the current quarters while keeping on the current administration system. Boundaries of LAU 2 regions are shown in Figure 5.17.

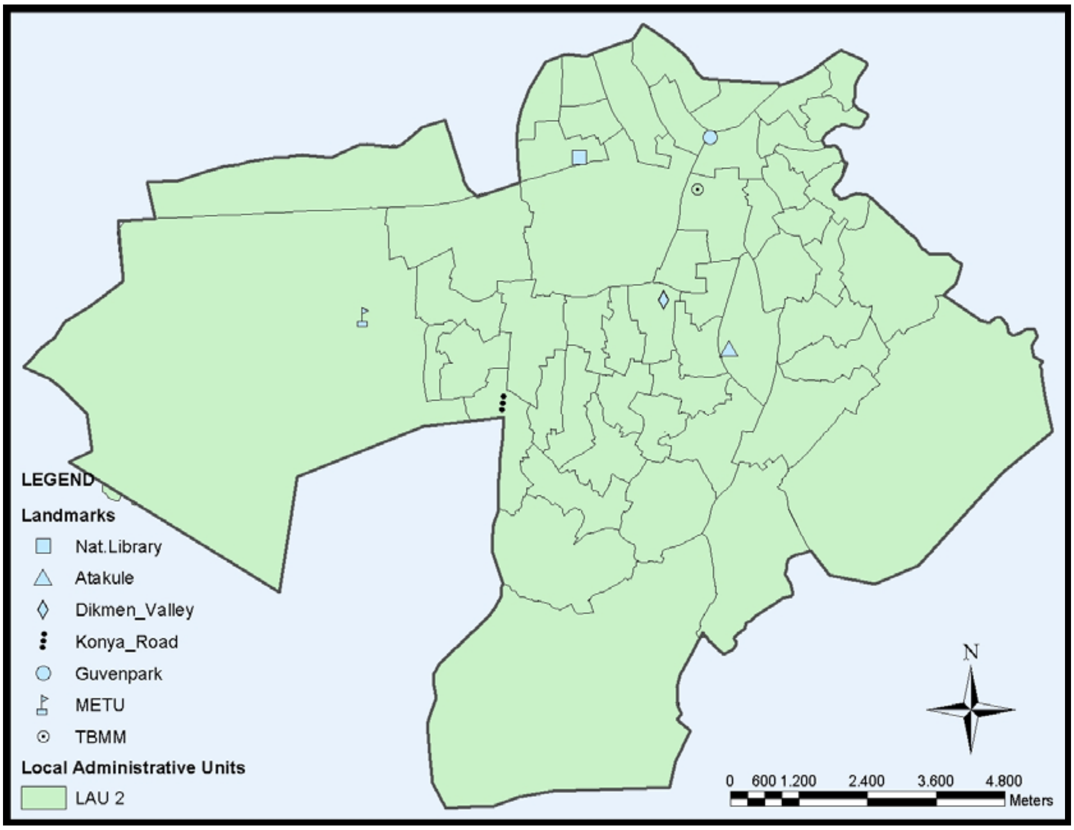


Figure 5.17 Boundaries of LAU 2 regions (Quarters) (54 units)

LAU 2, LAU 3 and LAU 4 regions are shown all together in Figure 5.18. In addition, Figure 5.19 shows a diagrammatic approach to the standard census geography hierarchy.

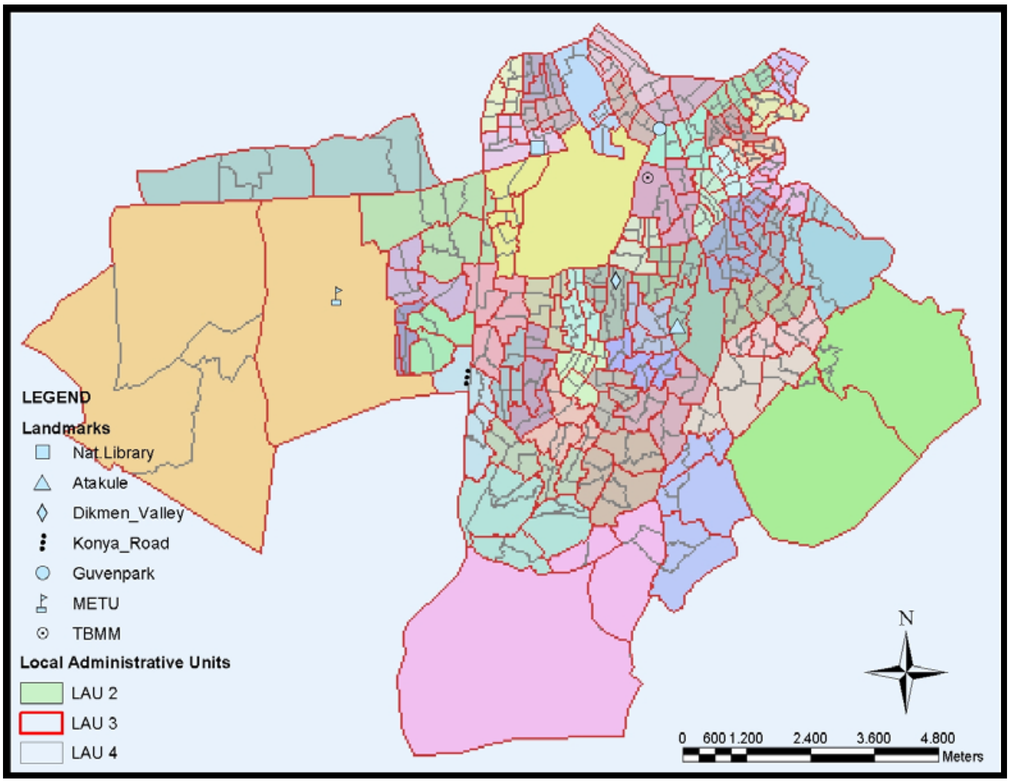


Figure 5.18 LAU 2, LAU 3 and LAU 4 hierarchy

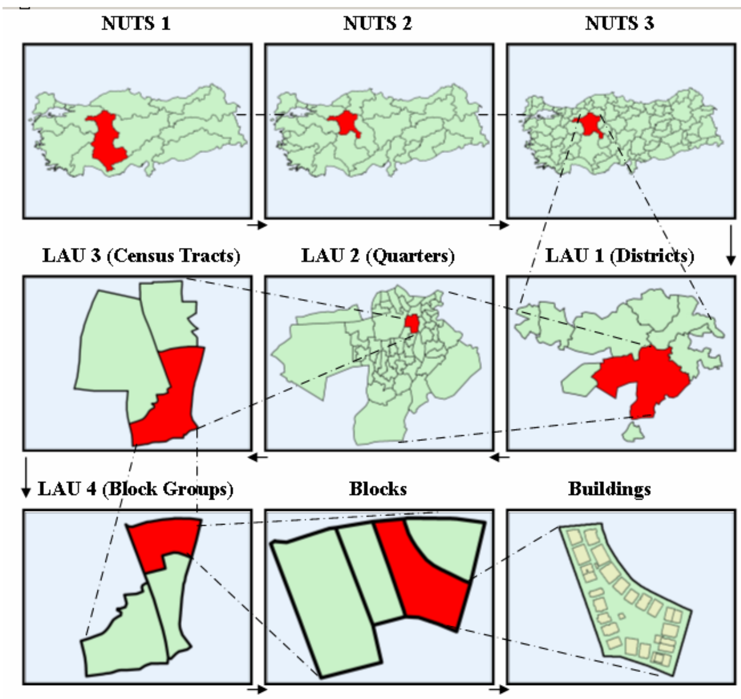


Figure 5.19 Standard Census Geography Hierarchy – An Overall Diagrammatic Approach

### 5.3.1 Region Coding

The proposed 2010 Census geography for Turkey consists of a very straightforward hierarchy, where lower level regions directly nest into higher level regions. The codes used to identify the regions must also be hierarchical:

12 units of NUTS 1 regions are coded as TR1, TR2, TR3, ...., TR9, TRA, TRB, and TRC.

26 units of NUTS 2 regions are coded as TR10, TR21, TR22, ...., TRA1, TRA2, ...TRC2, and TRC3 which are the lower levels of NUTS 1 codes.

81 units of NUTS 3 regions (provinces) are coded as TR100, TR211, TR212,….TRC32, TRC33, TRC34.

In the rest of the study, further codes will be assigned for only the case study area, not for all of Turkey; and these codes will be different than the NUTS level codes to prevent too long code expressions. There are 8 units of LAU 1 regions (districts) in Ankara, and these districts already have numerical expressions dedicated to them. The district code of Çankaya has been determined as “02” by the governmental organizations, so this coding has not been changed in the context of this thesis. Then the 2-digit plate number of the city, here ‘06’ for Ankara, is set prior to the 2-digit district code. As result, the region code of Çankaya is “0602” according to the census geography hierarchy.

There are 54 units of LAU 2 regions (quarters) defined inside of Çankaya. These areas are coded by two-digit letters like AA, AB, AC, …, ZY, and ZZ. Letters are preferred instead of numbers for this level because the code of this level comes after a four-digit number and two-digit letter will make it easier to remember the code. The letters used in this level of coding do not contain specific Turkish characters like Ç, Ğ, İ, Ö, Ş, and Ü. Thus, the total number of variations of this two-letter coding is equal to 23 x 23, which is equal to 529. This number is more than adequate for all of the possible applications in Turkey. For instance a LAU 2 region may be coded as “0602AA”, “0602AB”, “0602AC” etc.

There are 156 units of LAU 3 regions (census tracts) in Çankaya, but this does not mean that it is necessary to give 156 different numbers for each of them. Each LAU 2 region already has a unique code, and can contain a maximum of 9 units of LAU 3 regions because of the population limits. Therefore, giving a one-digit number after LAU 2 code is adequate to make any LAU 3 region unique. For instance a LAU 3 region may be coded as “0602AA1”, “0602AA2”, “0602AA3”, “0602AB1” etc.

There are 460 units of LAU 4 regions (block groups) in the case study area, but these regions do not need to be coded by 460 different numbers, as it was the case in LAU 3. Each LAU 3 region already has a unique code, and can contain a maximum of 9 units of LAU 4 regions because of the population limits. Therefore, giving a one-digit number after LAU 3 code is adequate to make any LAU 4 region unique. For instance a LAU 4 region may be coded as “0602AA11”, “0602AA12”, …, “0602AA21”, “0602AA22”, … etc. The hierarchy of region codes is summarized in Table 5.5;

Table 5.5 Hierarchy of existing and proposed regional codes;

| **Region Name** | **No. of Areas** | **Character** | **Zone Code** |
|---|---|---|---|
| **TURKEY** | 1 | 2 letters | TR |
| NUTS 1 | 12 | 1 letter/number | TR 5 |
| NUTS 2 | 26 | 1 number | TR 51 |
| NUTS 3 ( provinces ) | 81 | 2 numbers | TR510 |
| LAU 1 (districts, Ankara) | 8 | Plate No. + 2 no. | 0602 |
| LAU 2 (quarters, Çankaya) | 54 | 2 letters | 0602AA |
| LAU 3 (census tracts, Çankaya) | 156 | 1 number | 0602AA1 |
| LAU 4 ( block groups, Çankaya) | 460 | 1 number | 0602AA11 |

# CHAPTER 6

# CONCLUSIONS & RECOMMENDATIONS

This thesis aimed to create a new 'census geography' for Turkey based on the concept of 'small area statistics'. The general characteristics, application rules and methodologies of this census geography have been given in the previous chapters. In this chapter, the conclusions of the thesis and recommendations for further studies are discussed.

As it has been mentioned through the thesis, there is already a continuing process in Turkey to upgrade the current statistical system. Some concrete steps have also been taken such as creating the regional scale NUTS 1-2-3 statistical units and preparing important reports and draft laws necessary for further actions through the adaptation process to the European Union. This thesis has a considerable place among the similar others because it takes the first concrete step for defined further actions of Turkey about its statistical system. It brings a new approach to census geography concept in Turkey and emphasizes the importance of local/small area knowledge in addition to the regional. By the new system, it will be possible to make statistical comparisons with international countries, this will bring globally better harmonization of statistics, produced statistics will be more meaningful in global scale, and production and protection of time-series data will be possible.

In addition, policies may be monitored and assessed more accurately; because local and regional policies, and also the budgets, of all data users need to be allocated accurately. Therefore it is very important to target the resources effectively and efficiently to the related areas. In addition to the policy makers, local inhabitants also want to know more about the area where they live, and how it compares with other areas. Current census and survey statistics do not contain detailed information about local areas and do not give information about all the descriptive indicators.

Moreover, the kinds of current statistical geographic units are insufficient for serving different people from different organizations and professions, especially working in local scale; because the new geographic units will include place-specific information. This will make good contributions to many projects dealing with the local. The hierarchical structure of the proposed census geography includes varied sizes of statistical units; and this feature makes it flexible and more easily usable by many people.

As a result of the information based on small areas, the pinpointing of events will be easier and interactions among factors like unemployment, health, safety and education will be better understood at the local area level. For example, police forces will serve the community better and increase safety level using better geographic information to use in their geographic crime analyses. In addition, allocation of many public services may be targeted more precisely and their operational benefits will be increased, too. For instance; locations of clinics, schools, police stations, fire stations etc. will be determined after examining the local needs. Behind the service organizations, private companies may need these local data, too. For example, a branch of an insurance company has to be located on an area, where most people have money and no insurance.

The new system may also be used for better integration and use of existing nation-wide information held by private and public organizations. If all of separate data from different sources are combined on the same geographic basis and the necessary data are collected according to the new standards, nation-wide overlapping information will be eliminated. By a new common database, all the information may be brought together in a database management system, and for instance, it may be possible to relate different kinds of data from different sources. For example, a small statistical area may be identified in the new census geography as having high population density, overcrowded dwelling units also have children with low educational levels (data from Turkey Ministry of Education), poor health conditions (data from The Ministry of Health), and high unemployment ratio (data from State Institute of Statistics).

The proposed census geography also encourages interventions at local areas and based on small statistical areas. Individually these interventions may be looked small, but when they come together, the overall impact will be very large and benefits will be higher than costs. The availability of the new census geography will increase the efficiency of important private and public investments either by having the same cost and increasing the benefits or having the same benefits and decreasing the cost.

In addition to these powerful structure and possible benefits of the new census geography, there are some deficiencies of the proposed system and some further steps are needed for a better and nationwide application.

For instance, the major point of the methodology applied in the thesis is firstly to define as homogeneous as possible areas according to demographic and socio-economic characteristics of the inhabitants, and then to draw the boundaries of the new census geography units, each of which have specific population thresholds, from this point of view. Here, homogeneous areas have been defined based on 10 indicators; which are Literacy Ratio, University and Primary School Graduate Ratios, Proximity to Police Stations-Hospitals-Fire Stations, Average Household Size, Population Density, Unemployment Ratio, and Young Age Employment Ratio. These indicators and data about them have been relatively sufficient to realize the main aims of the thesis, and to create a new census geography. Nevertheless, it is necessary to increase the number of these indicators to have more accurate results. The additional indicators of ‘quality of life’ issues should be considered, including history, culture, ethnicity, geography, health, physical environment, crime, noise pollution, geology, air quality, aesthetics etc. However this thesis deals with local areas and political effects are not as effective as in the larger sclaes; it is necessary to include some indicators reflecting urban policy pattern in the different parts of the case study area.

Besides, through the decision making process only two multi-criteria decision making (MCDM) methods, Simple Additive Weighting (SAW) and Analytic Hierarchy Process (AHP) have been tested. Finally, all the decisions have been taken according to AHP due to unsatisfactory results of SAW. However, there are some other MCDM methods such as Value/Utility Function Approaches, Ideal Point Methods, Concordance Methods and Fuzzy Aggregation Operations etc. AHP has been preferred for the necessary operations because of its world-wide popularity and relatively easy application procedure. Nevertheless, it might be useful to apply the same procedure with all of the decision making methods to see which one gives the best results.

As different from an MCDM method, which has been used to find the homogeneous areas to the possible extent, some other methodologies may also be used for this purpose. For example, Exploratory Data Analysis (EDA) techniques may be more suitable in such a case, and especially in larger scale applications including much more larger datasets. EDA is defined in Wikipedia (2005) as "…that part of statistical practice concerned with reviewing, communicating and using data where there is a low level of knowledge about its cause system. Objectives of EDA are; to suggest hypotheses about the causes of observed phenomena, to assess assumptions on which statistical inference will be based, to support the selection of appropriate statistical tools and techniques, and to provide a basis for further data collection through surveys or experiments."

There may be many problematic areas where these proposals may become insufficient, however it is tried to create a nationwide applicable census geography by the already defined rules and methodologies. For instance, the new census geography has been created according to the de-facto population of the quarters; but there are specific areas such as city centers, central business districts etc. where the daily de-jure population is much higher than the de-facto. The average de-jure population of these areas may be determined and the boundaries of created statistical units may be looked over again. Another problematic subject is the ‘special places’, which are defined as block group when their population exceeds 300. In this study, only very big military areas and university campuses have been evaluated due to data unavailability, but it is necessary to search the locations and population numbers of all kinds of special places for a better classification.

In addition, rural areas of Turkey have a number of distinct characteristics that make it different from other countries. Turkey has a highly dispersed and differentiated rural settlement pattern with different characteristics and very low population numbers, which may not be adequate to satisfy the determined minimum population thresholds. Therefore, new rules and methodologies have to be created and tested for these areas. Not only the rural areas but also the seasonal settlements like secondary housing settlements, tourist areas etc. have to be evaluated again in another research project because of their changeable population characteristics.

Another important subject is the application of the new census geography methodology in squatter settlements. They are defined as temporary statistical units in this study and temporary boundaries are drawn for them. It is also proposed to redraw their boundaries in the direction of their development, and finally, converting some of them, which are completely rebuilt, to permanent statistical units. Nevertheless, the practical application may not be easy. The boundaries of squatter settlements may not be always drawn with discrete lines, and some areas may contain one within the other and mixed urban patterns. Therefore, while determining the temporary areas and redrawing the boundaries, it is very important to precisely separate the squatter settlements from the other urban areas.

Population thresholds of the small statistical units are very important to make the comparability possible. Therefore, population and housing pattern changes should not only be controlled in decennial census periods but also be updated in specific time periods. Through these updates, the integration of GIS and Remote Sensing technologies (including air and satellite photos etc.) will make important contributions. The efficient use of new information technologies will decrease both time and money consumptions. In addition, local redevelopment and etc. projects must be continuously followed up to make the necessary updates on the created statistical units such as converting the temporary statistical units to the permanent ones. Methodologically, the thesis proposes to split the statistical units exceeding the maximum population threshold, and to combine the statistical units having less population than the minimum population threshold to the next ones. These operations, of course, will cause changes on the drawn boundaries through years and seem to be opposite to the stability principle of statistical units. The number of these changes must be hold at minimum to protect the reliability of time-series data. Therefore, the areas which have important potential to increase or decrease their population numbers, must be previously defined as temporary statistical units; and they must be converted to the permanent ones when they are mostly complete their urbanization process.

For developing a nationwide census geography and producing decisions for a nation-wide application, the opinions of statistical data users such as governmental and private organizations, universities, institutions etc. are very important. A general consensus among them must be assured for a healthier and more useful census geography. Of course, through the thesis process, so many readings and comprehensive surveys have been made to understand their needs and estimate their opinions about a new census geography, but face–to-face interviews could give different results. Opinions of data users may increase both number of quality of life indicators and diversity of small statistical areas. In this way, the statistical data users may be able to find out all kinds of statistics about the related local areas.

After a new census geography has been created, it is necessary to develop a strong background, effective management and continuous upgrade in the system. The State Institute of Statistics (SIS) must develop and manage spatial databases; and these databases must be accessible for legislative purposes. SIS should also provide the necessary hardware, software and educated staff for the further updating procedures. In addition, it must "…continuously monitor the availability of spatial data potentially useful for both public and private purposes; represent the legislative interests in the broader community; and provide educational and outreach support for the legislative community" (Minnesota GIServices, 2005). Also, various kinds of representation techniques and map scales are needed for the different levels of the census geography hierarchy. Then, it is also necessary to adopt the collected data into these scales. Thus, the suitable techniques, scales, and generalization degrees of data for each statistical unit must be determined by the experts of the SIS. Only in this way can the new system be accurately applied and be successful.

This new system must also be accessible by all statistical data users without infringing the principle of personal confidentiality. The solution may be creation of a new 'web-based GIS system', which has all the levels of the hierarchical census geography integrated with the real-world features like streets, streamlines, mountains, railroads etc. and the related non-graphical data coming from censuses and surveys. Then people should have automated access to this database over the internet and retrieve necessary data, information and statistics about the area to which they are concerned as free. In addition to raw graphical and non-graphical data, users should see simple thematic maps of the selected indicators like population density, average household size, unemployment rates etc. to view graphical representation of data. Furthermore, through the further steps, this web-based GIS system may be more detailed and may also related with the other local, regional and national databases such as MERNIS (Central Population Management System) and TAKBIS (Turkish Land Registry and Cadastre Information System).

MERNIS is an "...e-government project which transfers all the identification information to electronic environment and it permits immediate updates and changes made in the identification information from 923 district centers in a secured way."

(Ministry of Internal Affairs, 2005) TAKBIS is also a strategic e-government project which has been created "…to improve land registration and cadastre data and operations. This project shall result in implementation of a Land Information System to improve operations of the land registration and cadastre offices, and to supply appropriate land related information to all its users, both private and public, municipalities in particular" (TUFUAB, 2004). If a central population database is built, public services are modernized, all the population information is transferred to digital environment, and finally a national network is established; the census geography will also serve as the geographic base for all of them and also be useful for the development of e-government concept in Turkey.

It is also definitely necessary make a feasibility study of such a nation-wide project before the application procedure of the new census geography proposal. Some districts may be selected as the pilot areas; and in 2010 population census, the proposed system may be applied. Firstly, accurate block structures are constructed for each pilot area, then LAU 4 statistical units are created by combining the blocks according to the population requirements, and then all the other levels are created from their lower level units in a nesting relationship. Next, through the census and post-census period, it is examined how the new system works, how much it is feasible, and how efficient to satisfy the needs of the statistical data users. Then, if necessary, some kind of arrangements and corrections may be made to develop the proposed census geography.

In conclusion, the new census geography proposal of this study is only a small (but important) contributory step for the already upgrading Turkish statistical system. If the created rules and methodologies are carefully examined and its deficiencies are completed through a widespread consensus to be reached through a public discussion, Turkey may have a much better census infrastructure than today for the 2010 Population Census and afterwards.

## REFERENCES

ABEL, A. Fernandez., (1996) "Expert Choice", ORMS – Online Edition, Software Review, Volume 23, Number 4, August 1996.
(http://www.lionhrtpub.com/orms/orms-8-96/software.html, visited on 05.02.2005)

ADB (Asian Development Bank), (2005) "Economic Analysis of Projects – Glossary", (http://www.adb.org/Documents/Guidelines/Eco_Analysis/glossary.asp, visited on 04.27.2005)

AHUJA, Brig J. S., (1999) "Census 2000: The GIS Way", 1999
(http://www.gisdevelopment.net/application/miscellaneous/misc024.htm, visited on 03.28.2005)

BACKER, Lars H., LUODE, Marja Tammilehto., GUIBLIN, Philippe., (2001) "In Search of a System of Small Statistical Areas", The Tandem Consortium, Background paper for the Workshop on Point-based and Area-based statistics, Work Session on Methodological Issues Involving the Integration of Statistics and Geography, 25 – 28 September 2001, Tallinn, Estonia.
(http://www.unece.org/stats/documents/2001/09/gis/crp.1.e.pdf, visited on 05.03.2005)

BACKER, Lars H., LUODE, Marja Tammilehto., GUIBLIN, Philippe., (2002) "A (feasibility) study towards a common geographical base for statistics across the European Union", Tandem_GIS I,
(http://epp.eurostat.cec.eu.int/cache/ITY_OFFPUB/KS-AN-02-001/EN/KS-AN-02-001-EN.PDF, visited on 01.26.2005)

Çankaya Municipality, (2005) "Çankaya with Numbers"
(http://www.cankaya-bld.gov.tr/cankaya_rakamlar.asp, visited on 03.15.2005)

CHRISMAN, Nicholas., (1996) "Exploring Geographic Information Systems", University of Washington, John Wiley & Sons, Inc. publication, 1996

DEMIR, Ömer. and TOPRAK, A.Ömer., (2004) “Turkish Statistical System: Current Situation and New Challenges”, United Nations Economic and Social Commission for Asia and Pasific Subcommittee on Statistics, Bangkok, Thailand, 18-20 February 2004 (http://www.unescap.org/stat/sos1/sos1_turkey.pdf, visited on 03.15.2005)

ERTUĞAY, Kıvanç., (2003) “Measurement and Evaluation of Fire Service Accessibility by Using GIS: A Case Study on Çankaya District, Ankara”, a Master of Science thesis submitted to the department of Geodetic and Geographic Information Technologies in graduate school of Natural and Applied Sciences of the Middle East Technical University, April 2003.

EDMONSTON, Barry., (1999) “Challenges For The U.S. Census In The Information Age”, June 1999.
(http://usinfo.state.gov/journals/itsv/0699/ijse/census.htm, visited on 04.05.2005)

EUROSTAT, (2004) “European Regional Statistics – Reference Guide”
(http://europa.eu.int/comm/eurostat/newcronos/reference/sdds/en/regio/regio_ref_guide.pdf, visited on 03.24.2005)

Fiji Bureau of Statistics, (2000) “Utilization of Census Data: GIS Modelling”, 2000
(http://www.ancsdaap.org/cencon2002/papers/Fiji/Fiji.PDF, visited on 03.24.2005)

GRANDA, Peter., (2003) “Census 2000: Accessing and Utilizing Data Resources”, Inter University Consortium for Political and Social Research (ICPSR), 2003
(http://www.icpsr.umich.edu/CENSUS2000/TRAINING/2001, visited on 12.22.2004)

Highlandgeographic, (2005) “Why GIS?”
(http://www.highlandgeographic.com/gis/ , visited on 04.15.2005)

HUT (Helsinki University of Technology), (2002) “Value Tree Analysis – THEORY”,
(http://www.mcda.hut.fi/value_tree/theory/theory.pdf, visited on 05.02.2005)

IRWIN, Elena., (2001) “Using Geographic Information Systems (GIS) in Economic Development: An Introduction to Applications, Terminology, and Software”, Ohio State University Extension, December 2001.
(http://aede.osu.edu/programs/ComRegEcon/gis/gisintro.htm, visited on 03.26.2005)

KONG, Yunfeng. and LIN, Hui., (1999) "Geographic Information Systems for Regional Business and Economy: A Theoretical Framework And System Analysis", The Chinese University of Hong Kong, China, June 1999, (www.umich.edu/~iinet/chinadata/geoim99/Proceedings/Kong_yunfeng.pdf, visited on 03.27.2005)

LAKHSMI, A. Sai Venkata., KUMAR, P. S. Bhavani. and BABA, J.Sai., (2001) "Embedded GIS for Census Mapping", India, June 2001, (http://www.gisdevelopment.net/application/miscellaneous/misc003.htm, 03.20.2005)

MALCZEWSKI, Jacek., (1999) "GIS and Multicriteria Decision Analysis", Department of Geography, University of Western Ontario, John Wiley & Sons Publication, 1999

MARTIN, David., 2001 "Geography for the 2001 Census in England" (http://www.statistics.gov.uk/geography/downloads/georoadshowpaper.pdf, 03.24.2005)

Ministry of Internal Affairs, (2005) "MERNIS - Central Population Management System" (http://www.nvi.gov.tr/11,Ana_Sayfa_Mernis_Ingilizce.html, visited on 04.04.2005)

Minnesota GIServices, (2005) "About LCC-GIS", (http://www.commissions.leg.state.mn.us/gis/html/aboutgis.html, visited on 03.20.2005)

NZCC (New Zealand City Councils) (2003), "Quality of Life Project - Qualification Levels", 10.07.2003 (http://www.bigcities.govt.nz/pdf/Qualifications.pdf, visited on 05.11.2005 )

ONS (Office for National Statistics), (2005a) "Geography and National Statistics", (http://www.statistics.gov.uk/geography/default.asp, visited on 03.28.2005)

ONS, (2005b) "Geographic Products" (http://www.statistics.gov.uk/geography/geog_products.asp, visited 03.28.2005)

ONS, (2005c) "Census Geography" (http://www.statistics.gov.uk/geography/census_geog.asp, visited on 03.28.2005)

ONS, (2005d) “Super Output Areas (SOAs)”
(http://www.statistics.gov.uk/geography/soa.asp, visited on 03.28.2005)

RATCLIFFE, J. H., (1999) “The Ecological Fallacy”
(http://www.jratcliffe.net/research/ecolfallacy.htm, visited on 03.13.2005)

SACTAQC (The Sacramento Transportation and Air Quality Collaborative), (2005) “A Glossary Of Terms - Land Use Planning”
(www.sactaqc.org/Resources/primers/Glossary_Land_Use.htm, visited on 05.11.2005)

SARAÇGİL, Ayse., (1999) "The Gecekondu and Turkish Modernity" Environmental Design: Journal of the Islamic Environmental Design Research Centre 1-2 (1999),
(http://archnet.org/library/sites/one-site.tcl?site_id=3941, visited on 03.15.2005)

SASWATI, Paik., (2001) “GIS: Bringing Sense to Census”, Jawaharlal Nehru University, New Delhi, June 2001
(http://www.gisdevelopment.net/magazine/gisdev/2001/june/gisbsc.shtml, 03.20.2005)

SINGH, R.P., (2001) “Application of GIS in Census Of India 2001”, Office of the Registrar General of India,
(http://incaindia.org/applicatioofGIS.pdf, visited on 03.20.2005)

SIS (State Institute of Statistics), (2002a) “2002 Progress Report, Chapter 12: Statistics”
(www.die.gov.tr/abkd/English_ab/2002istIlerleme_eng.htm, visited on 03.17.2005)

SIS, (2002b) “Country Paper: Republic Of Turkey”, Economic and Social Commission for Asia and the Pacific (ESCAP) Committee on Statistics - Thirteenth Session, Bangkok, Thailand, November 2002,
(http://www.unescap.org/stat/cos13/cos13_turkey.pdf, visited on 03.25.2005)

SIS, (2004) “Press Release: Results of 2002 Poverty Study”, 04.13.2004,
(www.die.gov.tr/english/SONIST/YOKSL/200404/200404.htm, visited on 05.10.2005)

SIS, (2005) “Data Dictionary”
(http://www.die.gov.tr/TURKISH/SOZLUK/dataa.html, visited on 03.22.2005)

SIS, (2005a) "Statistics history of Turkey"
(http://www.die.gov.tr/English/history.html, visited on 3.18.2005)

SIS, (2005b) "Importance of SIS"
(http://www.die.gov.tr/onemIslev.htm, visited on 3.18.2005)

TBMM, (2004) "Municipality Law", Law No: 5272, Accept Date: 07.12.2004
(http://www.tbmm.gov.tr/kanunlar/k5272.html, visited on 03.25.2005)

TOIDA, Koki., (1999) "Use Of GIS for - Census Mapping In Japan", Government of Japan - Statistics Bureau, Management and Coordination Agency, July 1999,

TUFUAB (Turkish National Society for Photogrammetry and Remote Sensing), (2004) "National Report Of Turkey On Photogrammetry and Remote Sensing for 2000 – 2004", (www.isprs.org/istanbul2004/comm6/papers/720.pdf, visited on 03.20.2005)

UNDP (United Nations Development Programme), (2004) "HUMAN DEVELOPMENT REPORT 2004 - Cultural Liberty in Today's Diverse World"
(http://hdr.undp.org/reports/global/2004/pdf/hdr04_complete.pdf, visited on 05.11.2005)

UN Statistics Division, (2004) "Integration of GPS, Digital Imagery and GIS with Census Mapping", UNITED NATIONS SECRETARIAT, Department of Economic and Social Affairs - Statistics Division, New York, 15-17 September 2004,

US (United States) Census Bureau, (1994a) "Geographic Areas Reference Manual - Census Bureau Geography", Geography Division of US Census Bureau, 1994,
(http://www.census.gov/geo/www/GARM/Ch1GARM.pdf, visited on 03.08.2005)

U.S. Census Bureau, (1994b) "Geographic Areas Reference Manual – Geographic Overview", Geography Division of US Census Bureau, 1994,
(http://www.census.gov/geo/www/GARM/Ch2GARM.pdf, visited on 03.08.2005)

U.S. Census Bureau, (1994c) "Geographic Areas Reference Manual – Census Blocks & Block Groups", Geography Division of US Census Bureau, 1994,
(http://www.census.gov/geo/www/GARM/Ch11GARM.pdf, visited on 03.08.2005)

U.S. Census Bureau, (2000) "Participant Statistical Areas Program Guidelines - 2000" (http://www.census.gov/geo/www/statarea.pdf, visited on 03.08.2005)

U.S. Census Bureau, (2002) "Glossary", (http://www.census.gov/geo/www/tiger/glossary.html#glossary, visited on 02.25.2005)

U.S. Census Bureau, (2004) "TIGER Overview", (http://www.census.gov/geo/www/tiger/overview.html, visited on 04.04.2005)

UN Statistics Division, (2000) "Handbook on GIS and Digital Mapping for Population and Housing Census Country Report", Department of Economic and Social Affairs, Member Countries of the Working Party, ESCAP

VIRDEE, Dev., (2001) "Strategies And Approaches For Small Area Statistics", Invited paper prepared by the Office for National Statistics, Conference of European Statisticians, 49th plenary session, Geneva, June 2001. (www.unece.org/stats/documents/ces/2001/19.e.pdf,visited on 05.03.2005)

WebGIS, (2005) "Glossary – General Terms" (http://www.webgis.net/cms.php/glossary.html, visited on 04.15.2005)

Wikipedia, (2005) "Exploratory Data Analysis", Wikipedia the free encyclopedia, May, 2005. (http://en.wikipedia.org/wiki/Exploratory_data_analysis, visited on 06.15.2005)

World Bank, (2003) "Poverty Assessment Report – Turkey: Poverty and Coping after Crises", Volume One: Main Report, Report No. 24185, Human Development Unit Europe and Central Asia Region, July 2003. (http://lnweb18.worldbank.org/eca/ecspeExt.nsf/AttachedById/D5CA119674EC991E85256E76006F8313/$FILE/TurkeyPA_%20Vol%201.pdf, visited on 05.12.2005)

YETIK, Arzu., (2003) "The Use of Geographic Information Systems in Decision Support System & An Application Concerning Small Area Districting for Statistical Purposes in Zafer Sub-District of Istanbul-Bahcelievler", State Institute of Statistics – Expertness Thesis, Ankara, 2003.

## APPENDIX A

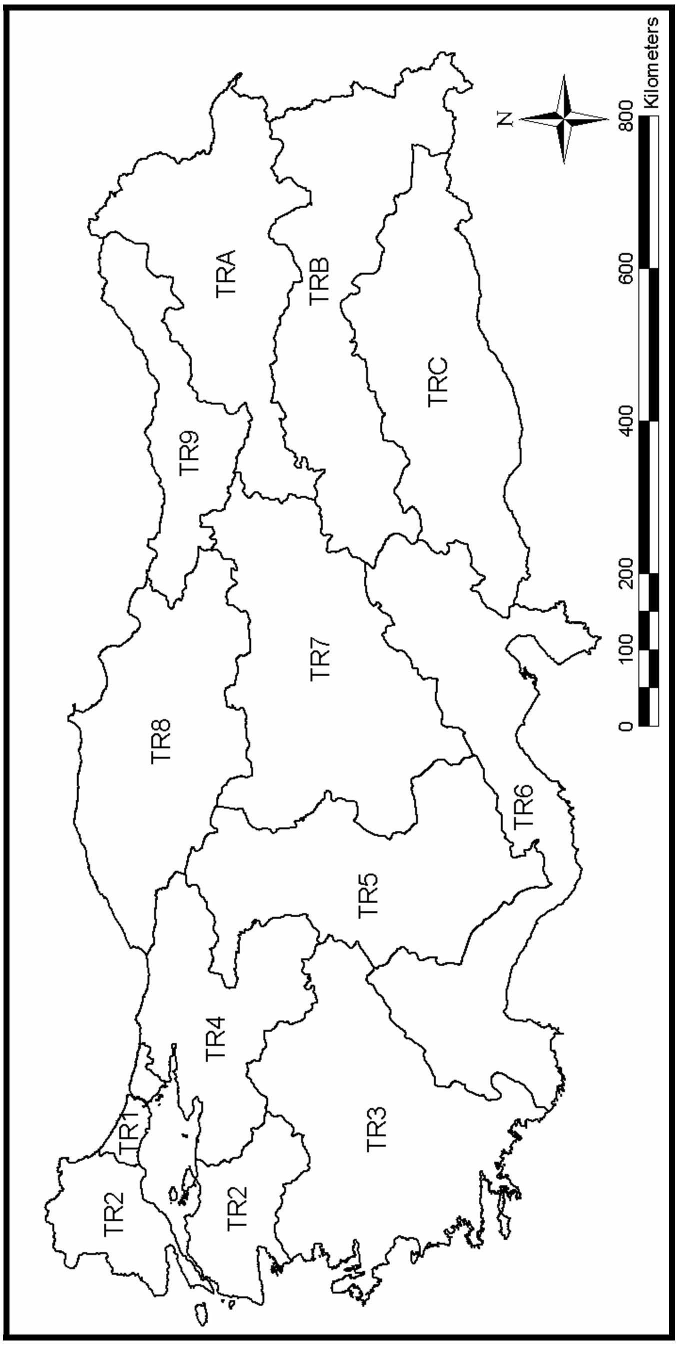


Figure A.1 Boundaries of NUTS Regions in Turkey – NUTS 1 (12 units)

**APPENDIX A (cntd.)**

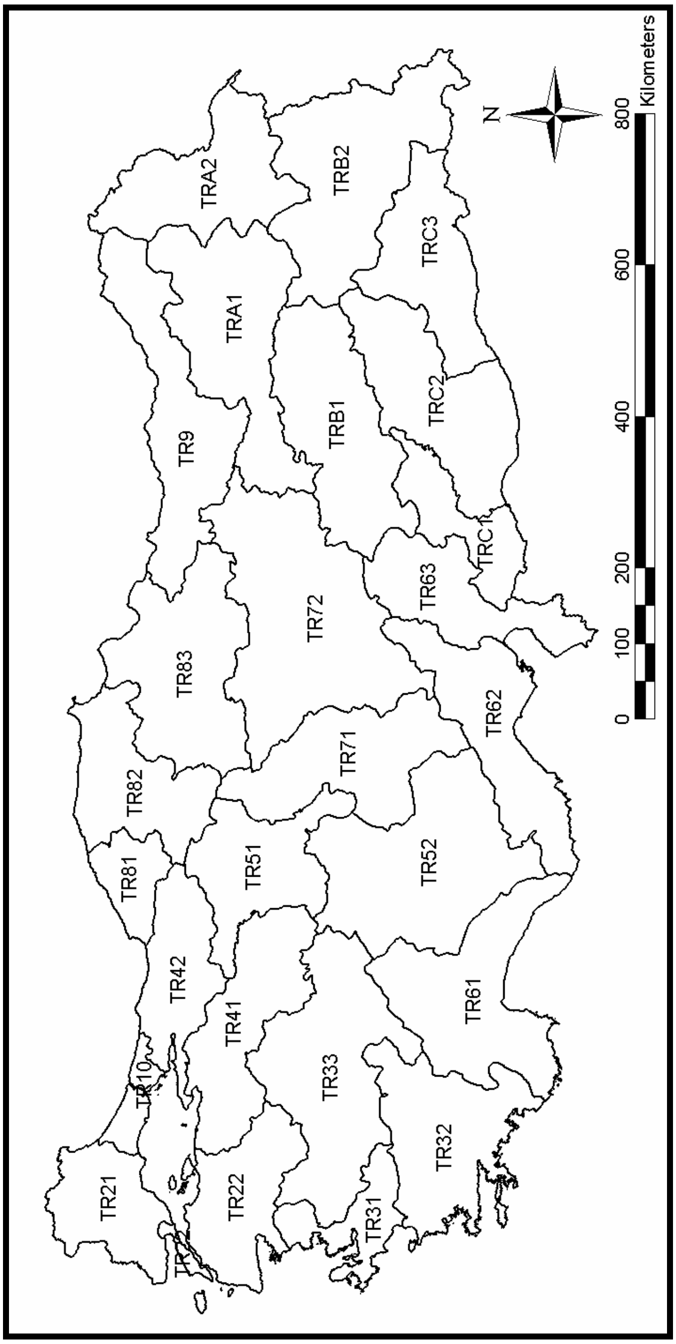


Figure A.2 Boundaries of NUTS Regions in Turkey – NUTS 2 (26 units)

**APPENDIX A (cntd.)**

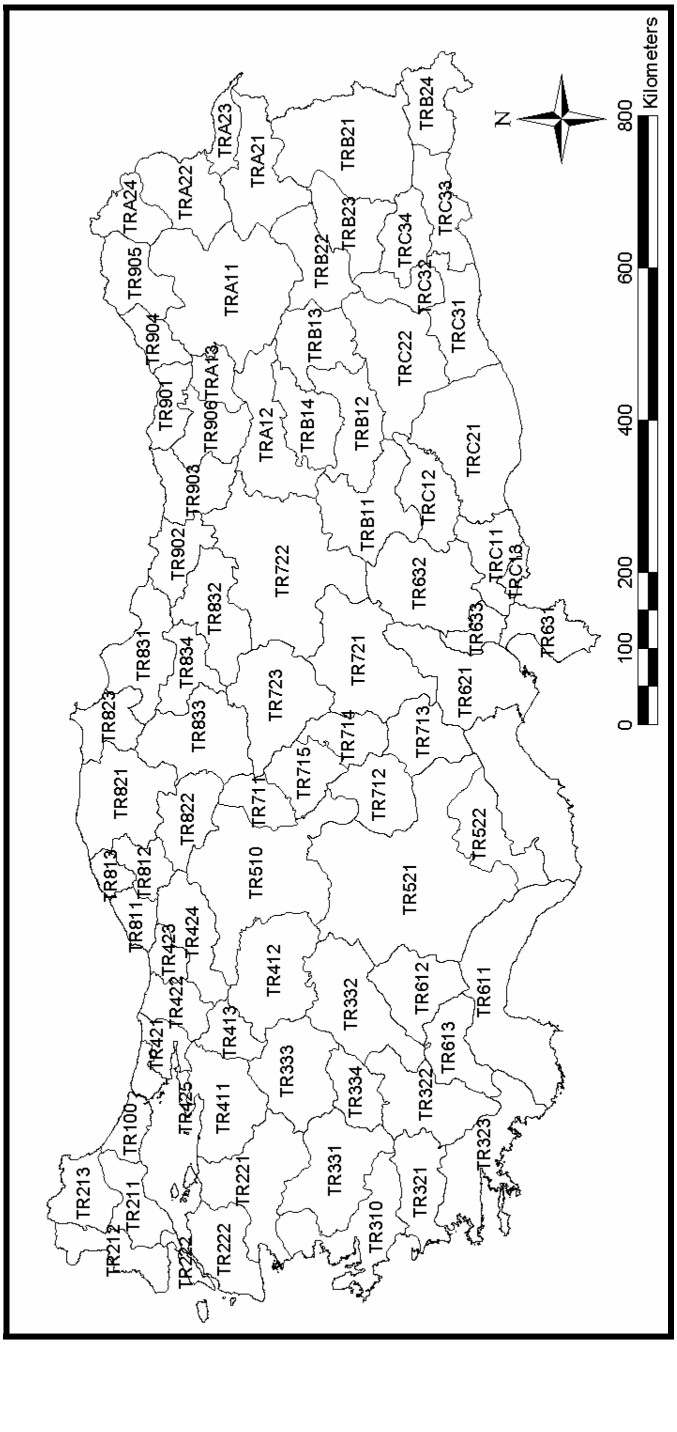


Figure A.3 Boundaries of NUTS Regions in Turkey – NUTS 3 (81 units)

# APPENDIX B

## United Nations - Human Development Indicators (2004)

<u>I. MONITORING HUMAN DEVELOPMENT: ENLARGING PEOPLE'S CHOICES</u>

1 Human development index

2 Human development index trends

3 Human and income poverty: developing countries

4 Human and income poverty: OECD, Central & Eastern Europe & CIS

<u>II. TO LEAD A LONG AND HEALTHY LIFE</u>

5 Demographic trends

6 Commitment to health: resources, access and services

7 Water, sanitation and nutritional status

8 Leading global health crises and risks

9 Survival: progress and setbacks

<u>III. TO ACQUIRE KNOWLEDGE</u>

10 Commitment to education: public spending

11 Literacy and enrolment

12 Technology: diffusion and creation

<u>IV. TO HAVE ACCESS TO THE RESOURCES NEEDED FOR A DECENT STANDARD OF LIVING</u>

13 Economic performance

14 Inequality in income or consumption

15 Structure of trade

## APPENDIX C

Table C.1 General Information about Çankaya Quarters

| QuarterName | Pop.# | Dwelling # | Hhold | Illiteracy | Univ. | Primry | Young | Unemp |
|---|---|---|---|---|---|---|---|---|
| Akpinar | 4983 | 1569 | 3.18 | 0.08 | 0.10 | 0.51 | 0.11 | 0.11 |
| Anittepe | 6396 | 2492 | 2.57 | 0.02 | 0.37 | 0.19 | 0.02 | 0.08 |
| Arkatopraklik | 3183 | 1352 | 2.35 | 0.06 | 0.13 | 0.39 | 0.08 | 0.17 |
| Asagi Dikmen | 9721 | 2547 | 3.82 | 0.01 | 0.40 | 0.16 | 0.02 | 0.05 |
| Asagi Imrahor | 165 | 63 | 2.62 | 0.04 | 0.02 | 0.71 | 0.19 | 0.02 |
| AsagiOvecler | 8516 | 2198 | 3.87 | 0.09 | 0.31 | 0.29 | 0.06 | 0.09 |
| Asikpasa | 4940 | 1699 | 2.91 | 0.08 | 0.07 | 0.54 | 0.14 | 0.12 |
| Ata | 5277 | 2064 | 2.56 | 0.05 | 0.11 | 0.52 | 0.10 | 0.12 |
| Aydinlar | 6690 | 2622 | 2.55 | 0.02 | 0.26 | 0.31 | 0.04 | 0.12 |
| Ayranci | 20525 | 8759 | 2.34 | 0.02 | 0.42 | 0.18 | 0.02 | 0.09 |
| Aziziye | 12769 | 6538 | 1.95 | 0.09 | 0.44 | 0.17 | 0.03 | 0.09 |
| Bademlidere | 1769 | 504 | 3.51 | 0.08 | 0.05 | 0.55 | 0.09 | 0.16 |
| Bagcilar | 1942 | 743 | 2.61 | 0.02 | 0.20 | 0.43 | 0.10 | 0.14 |
| Bahcelievler | 15807 | 5721 | 2.76 | 0.04 | 0.39 | 0.16 | 0.02 | 0.08 |
| Balgat | 6887 | 1741 | 3.96 | 0.02 | 0.19 | 0.31 | 0.05 | 0.06 |
| Barbaros | 7304 | 3921 | 1.86 | 0.05 | 0.42 | 0.17 | 0.04 | 0.08 |
| Bayraktar | 5349 | 1635 | 3.27 | 0.04 | 0.23 | 0.33 | 0.06 | 0.07 |
| Birlik | 22354 | 9028 | 2.48 | 0.10 | 0.35 | 0.24 | 0.06 | 0.10 |
| Boztepe | 2401 | 591 | 4.06 | 0.03 | 0.03 | 0.59 | 0.16 | 0.16 |
| Buyukesat | 9257 | 4733 | 1.96 | 0.03 | 0.47 | 0.17 | 0.03 | 0.08 |
| Camlitepe | 7653 | 3000 | 2.55 | 0.02 | 0.29 | 0.25 | 0.05 | 0.13 |
| Cankaya | 10890 | 3781 | 2.88 | 0.02 | 0.33 | 0.19 | 0.02 | 0.04 |
| Cebeci | 7446 | 2658 | 2.80 | 0.09 | 0.24 | 0.21 | 0.05 | 0.13 |
| Cengiz Karaca | 4470 | 1580 | 2.83 | 0.07 | 0.13 | 0.46 | 0.11 | 0.13 |
| Cevdet Ozdemir | 7088 | 2358 | 3.01 | 0.06 | 0.18 | 0.37 | 0.07 | 0.13 |
| Cevizlidere | 7095 | 2635 | 2.69 | 0.05 | 0.03 | 0.63 | 0.16 | 0.12 |
| Cukuranbar | 5271 | 3297 | 1.60 | 0.01 | 0.15 | 0.51 | 0.12 | 0.09 |
| Cumhuriyet | 766 | 294 | 2.29 | 0.03 | 0.14 | 0.40 | 0.07 | 0.04 |
| Devlet | 972 | 351 | 2.77 | 0.08 | 0.13 | 0.24 | 0.04 | 0.01 |
| Dilekler | 2130 | 743 | 2.87 | 0.03 | 0.05 | 0.59 | 0.14 | 0.15 |
| Dogus | 3614 | 1545 | 2.34 | 0.05 | 0.31 | 0.26 | 0.03 | 0.12 |
| Ehlibeyt | 3385 | 902 | 3.75 | 0.10 | 0.28 | 0.32 | 0.08 | 0.08 |

Table C.1 General Information about Çankaya Quarters (cntd.)

| QuarterName | Pop.# | Dwelling # | Hhold | Illiteracy | Univ. | Primry | Young | Unemp |
|---|---|---|---|---|---|---|---|---|
| Ellinci Yil | 2095 | 796 | 2.63 | 0.02 | 0.02 | 0.61 | 0.15 | 0.17 |
| Emek | 25262 | 9680 | 2.61 | 0.05 | 0.37 | 0.19 | 0.03 | 0.12 |
| Ertugrulgazi | 8582 | 3123 | 2.75 | 0.03 | 0.20 | 0.32 | 0.07 | 0.14 |
| Erzurum | 3342 | 1385 | 2.41 | 0.02 | 0.25 | 0.23 | 0.05 | 0.12 |
| Esatoglu | 4553 | 2057 | 2.21 | 0.02 | 0.38 | 0.20 | 0.03 | 0.11 |
| Eti | 1232 | 991 | 1.24 | 0.02 | 0.37 | 0.24 | 0.07 | 0.07 |
| Fakulteler | 6753 | 2841 | 2.38 | 0.02 | 0.30 | 0.21 | 0.03 | 0.12 |
| Fidanlik | 1866 | 880 | 2.12 | 0.02 | 0.37 | 0.15 | 0.05 | 0.12 |
| Gaziosmanpasa | 5038 | 2446 | 2.06 | 0.09 | 0.47 | 0.17 | 0.02 | 0.07 |
| Gokkusagi | 4322 | 1183 | 3.65 | 0.07 | 0.02 | 0.64 | 0.15 | 0.13 |
| Gokturk | 3310 | 1341 | 2.47 | 0.02 | 0.14 | 0.44 | 0.10 | 0.11 |
| Guven | 11360 | 5465 | 2.08 | 0.03 | 0.45 | 0.16 | 0.04 | 0.09 |
| Guzeltepe | 6090 | 3656 | 1.67 | 0.04 | 0.41 | 0.21 | 0.04 | 0.10 |
| Harbiye | 24334 | 8831 | 2.76 | 0.05 | 0.29 | 0.27 | 0.05 | 0.13 |
| Hilal | 3427 | 1753 | 1.95 | 0.05 | 0.20 | 0.41 | 0.11 | 0.09 |
| Huzur | 14309 | 5097 | 2.81 | 0.05 | 0.23 | 0.35 | 0.07 | 0.12 |
| Ileri | 5968 | 2283 | 2.61 | 0.05 | 0.20 | 0.32 | 0.06 | 0.14 |
| Ilkadim | 12839 | 6148 | 2.09 | 0.11 | 0.24 | 0.33 | 0.06 | 0.11 |
| Ilker | 4051 | 1247 | 3.25 | 0.05 | 0.05 | 0.56 | 0.13 | 0.16 |
| Incesu | 4522 | 1697 | 2.66 | 0.03 | 0.22 | 0.30 | 0.06 | 0.14 |
| Isci Bloklari | 20818 | 8687 | 2.40 | 0.03 | 0.31 | 0.24 | 0.03 | 0.11 |
| Karakusunlar | 12898 | 5239 | 2.46 | 0.08 | 0.41 | 0.21 | 0.05 | 0.10 |
| Karapinar | 2922 | 1118 | 2.61 | 0.02 | 0.06 | 0.56 | 0.15 | 0.12 |
| Kavaklidere | 9088 | 4114 | 2.21 | 0.02 | 0.41 | 0.16 | 0.03 | 0.08 |
| Kazim Ozalp | 6283 | 2508 | 2.51 | 0.07 | 0.47 | 0.18 | 0.02 | 0.07 |
| Keklikpinari | 10823 | 3852 | 2.81 | 0.08 | 0.26 | 0.36 | 0.07 | 0.12 |
| Kirkkonaklar | 9055 | 3283 | 2.76 | 0.02 | 0.14 | 0.47 | 0.11 | 0.14 |
| Kizilay | 2539 | 856 | 4.42 | 0.08 | 0.24 | 0.22 | 0.06 | 0.06 |
| Kizilirmak | 4036 | 1589 | 2.54 | 0.02 | 0.13 | 0.50 | 0.12 | 0.11 |
| Kocatepe | 1210 | 636 | 1.90 | 0.03 | 0.31 | 0.15 | 0.08 | 0.13 |
| Korkut Reis | 2140 | 856 | 3.52 | 0.03 | 0.25 | 0.25 | 0.05 | 0.10 |
| Kucukesat | 3711 | 1419 | 2.62 | 0.02 | 0.34 | 0.17 | 0.02 | 0.13 |
| Kültür | 4526 | 2312 | 1.96 | 0.02 | 0.38 | 0.19 | 0.03 | 0.10 |
| Malazgirt | 4012 | 1676 | 2.39 | 0.08 | 0.17 | 0.45 | 0.08 | 0.13 |
| Maltepe | 10999 | 4610 | 2.39 | 0.02 | 0.36 | 0.18 | 0.04 | 0.10 |
| Mebusevleri - Anitkabir | 5639 | 2094 | 2.69 | 0.02 | 0.39 | 0.16 | 0.03 | 0.07 |
| Mesrutiyet | 2142 | 839 | 1.55 | 0.03 | 0.27 | 0.21 | 0.08 | 0.09 |

Table C.1 General Information about Çankaya Quarters (cntd.)

| QuarterName | Pop.# | Dwelling # | Hhold | Illiteracy | Univ. | Primry | Young | Unemp |
|---|---|---|---|---|---|---|---|---|
| Metin Akkus | 4152 | 1295 | 3.21 | 0.06 | 0.20 | 0.37 | 0.07 | 0.13 |
| Metin Oktay | 5135 | 1555 | 3.30 | 0.07 | 0.11 | 0.44 | 0.08 | 0.12 |
| Mevlut Meric | 6690 | 2422 | 2.76 | 0.07 | 0.46 | 0.16 | 0.03 | 0.08 |
| Mimarsinan | 3273 | 1113 | 2.94 | 0.10 | 0.05 | 0.57 | 0.13 | 0.17 |
| Muhsin Ertugrul | 4211 | 1689 | 2.49 | 0.03 | 0.31 | 0.25 | 0.03 | 0.11 |
| Murat | 4463 | 1585 | 2.82 | 0.06 | 0.20 | 0.35 | 0.06 | 0.12 |
| Mursel Uluc | 12711 | 4037 | 3.15 | 0.07 | 0.15 | 0.43 | 0.09 | 0.12 |
| Mustafa Kemal | 5347 | 1742 | 3.07 | 0.03 | 0.43 | 0.23 | 0.04 | 0.08 |
| Naci Cakir | 10584 | 3536 | 2.99 | 0.05 | 0.20 | 0.35 | 0.05 | 0.12 |
| Namik Kemal | 1617 | 479 | 3.38 | 0.01 | 0.44 | 0.19 | 0.04 | 0.05 |
| Nasuh Akar | 4713 | 1634 | 2.88 | 0.04 | 0.35 | 0.25 | 0.04 | 0.09 |
| Oguzlar | 9482 | 3238 | 2.93 | 0.06 | 0.24 | 0.34 | 0.06 | 0.10 |
| On Cebeci | 6119 | 2513 | 2.43 | 0.03 | 0.34 | 0.21 | 0.05 | 0.12 |
| Oran | 11200 | 4411 | 2.54 | 0.02 | 0.48 | 0.16 | 0.03 | 0.06 |
| Orta Imrahor | 583 | 140 | 4.16 | 0.10 | 0.01 | 0.69 | 0.18 | 0.04 |
| Ortadogu | 19760 | 4267 | 4.63 | 0.01 | 0.25 | 0.11 | 0.07 | 0.09 |
| Osman Temiz | 8637 | 3261 | 2.65 | 0.04 | 0.27 | 0.30 | 0.05 | 0.12 |
| Ovecler | 8190 | 3441 | 2.38 | 0.06 | 0.16 | 0.42 | 0.08 | 0.11 |
| Remzi Oguz Arik | 6523 | 3558 | 1.83 | 0.02 | 0.45 | 0.17 | 0.03 | 0.10 |
| Saglik | 1024 | 486 | 1.70 | 0.05 | 0.28 | 0.26 | 0.04 | 0.17 |
| Sancak | 18137 | 6604 | 2.75 | 0.05 | 0.29 | 0.31 | 0.06 | 0.11 |
| Seyran | 9200 | 3828 | 2.40 | 0.05 | 0.23 | 0.30 | 0.05 | 0.13 |
| Sogutozu | 5079 | 1197 | 4.24 | 0.02 | 0.29 | 0.18 | 0.06 | 0.10 |
| Sokullu Mehmet Pasa | 10789 | 3911 | 2.76 | 0.05 | 0.23 | 0.34 | 0.06 | 0.11 |
| Tinaztepe | 7577 | 3533 | 2.14 | 0.03 | 0.32 | 0.23 | 0.03 | 0.12 |
| Topraklik | 1867 | 668 | 2.79 | 0.05 | 0.19 | 0.34 | 0.05 | 0.15 |
| Umut | 9121 | 2990 | 3.05 | 0.05 | 0.17 | 0.39 | 0.06 | 0.14 |
| Y.Bahçelievler | 19061 | 8277 | 2.30 | 0.09 | 0.02 | 0.77 | 0.03 | 0.11 |
| Yesilkent | 812 | 270 | 3.01 | 0.02 | 0.41 | 0.16 | 0.20 | 0.10 |
| Yildizevler | 11869 | 5163 | 2.30 | 0.04 | 0.29 | 0.30 | 0.06 | 0.11 |
| Yucetepe | 7171 | 2076 | 3.45 | 0.01 | 0.31 | 0.17 | 0.03 | 0.06 |
| Yukari Dikmen | 12652 | 3596 | 3.52 | 0.06 | 0.24 | 0.39 | 0.10 | 0.12 |
| Yukari Ovecler | 3045 | 1090 | 2.79 | 0.07 | 0.15 | 0.49 | 0.12 | 0.11 |
| Yuzuncu Yil | 7109 | 3091 | 2.30 | 0.02 | 0.41 | 0.20 | 0.05 | 0.09 |
| Zafertepe | 3988 | 1453 | 2.74 | 0.10 | 0.08 | 0.57 | 0.12 | 0.17 |