CHAPTER 6

A Geographical Information System used in the 2011 Census of Population and Housing

Maria Refalo, Silvan Zammit, Saviour Formosa and Ashley Hili

Introduction

The National Statistics Office is Malta's National Authority entrusted to carry out the census. The population statistic is only part of the whole exercise, thus the purpose of the census has evolved. The exercise has evolved over time with questions creating a spectrum of the Maltese society situation. The fact that it is carried out every 10 years, it allow for comparisons to be systematically arranged and within a time frame which is not short nor long. The census on its own is an extensive project to carry out, nevertheless the information gained from this national project provides data which needs to be turned into information, giving more value to the data gathered.

The national census being a vast data source, if integrated with GIS can produce several beneficial statistical outputs. The decennial count of residents, households and dwellings is a pool for statisticians to create progression, differences and a comparison between the different census years. Having GIS to analyse the data through its several tools can create a more systematic, efficient and clear results. The polygonisation process of the Enumeration Areas (EAs) was the next to be created, specifically for the purpose of the census statistics to be incorporated and thus can be spatially visualised. The process to create the polygons was made systematically so as to have a clearer way to visualise exactly how the plotting should be made. Plotting the EAs will not be a static map as it could be used further on with other census outputs, or if changes are necessary, amendments will be needed to be done.

Being able to integrate data from the census can be further on used to analyse the distribution of the questionnaire, giving a more representative output for policy makers. Furthermore, having the census with a spatial aspect was a must to relate to the contemporary needs.

The Population and Housing Census in Malta and the EU

The Census of Population and Housing provides an official count of residents, households, and dwellings in a country. It is normally a decennial exercise that allows for comparison of information over relatively fixed periods of time and presents a rich socioeconomic profile of the population as well as the dwelling stock characteristics, covering occupied and vacant property.

Every census exercise is specific and exhibits its own special characteristics. The last census in Malta was undertaken in 2011 and was the seventeenth of its kind since 1842. It was part of a European-union-wide round, since all Member States conducted their own national censuses. As in almost half of these countries, including Malta, the census is carried out the traditional way, whereby primary data is systematically collected by field officers (or enumerators) directly from individuals through questionnaires. Computer assisted and mixed data collection modes are gaining in popularity across the European Statistical System over time and some countries (particularly the Nordic) have a long tradition in undertaking fully register-based censuses. Mixed-mode data collection methods, involving a mix of different approaches, are also quite popular.

The census has a solid legal basis, underpinned by the provisions of the Maltese Census Act of 1948. Additionally, a number of Regulations of the European Parliament and of the Council establish common rules for the provision and dissemination of comprehensive census data in terms of methodologies and metadata.

The Census Questionnaire

The census provides a snapshot of a country's population at a particular point in time known as 'census night', which, in case of the 2011 census in Malta, corresponded to midnight of 20 November. This means that all information collected in the last census process had been recorded relative to this instance. This minimised problems of counting vacationers, newborns, seasonal labourers and other seasonal population in the country.

Census variables, concepts and definitions which are of general European interest are mandatory as stipulated by Eurostat's Regulations although countries are free to include any additional variables for national interest and use. In case of the 2011 round, (EC) No 763/2008 and (EC) No 1201/2009 applied. National issues were also addressed following a series of consultation meetings held with a number of organisations months before the official launch of the census, and the questionnaire was also piloted among a random sample of households for evaluation. In addition to aforementioned Regulations, census

questionnaires must also ensure continuity as well as conformity to the recommendations, concepts and definitions on population censuses of the United Nations Economic Commission for Europe (UNECE).

The same questionnaire was used to collect information from private and institutional households, and respondents were able to request either a Maltese or English-language version. It was split into two parts, with the first focusing on socioeconomic characteristics of the population and the second part dealing with housing.

Constructing Enumeration Areas

For operational issues, Malta was divided into six districts according to the Local Administrative Unit (LAU) classification which comprises all the 68 localities (equivalent to local councils) in the country. All localities were further divided into 1021 geographical units called Enumeration Areas. Each EA was composed of a number of streets, or parts thereof, in a particular locality whilst taking into consideration both the number of dwellings included in each EA as well as the geographical area covered.

The initial step for delineating EAs was to identify the count of private dwellings in each street in Malta. A list of unique streets was extracted from the 2011 government's corporate database known as the Common Database (CdB), which contains details about Maltese residents in possession of a Maltese identity card. Other auxiliary databases available in the country were also consulted to ensure that a final comprehensive list of streets is created together with an estimate of occupied and vacant property in each street. This list was then provided to the Malta Environment and Planning Authority (MEPA) who constructed the EAs and represented them as maps in electronic format.

Each map was inspected in detail for any possible overlap, gaps and other inconsistencies. Particular attention was given to streets which spanned over multiple EAs to ensure that boundaries were clearly and distinctly marked.

Census Operations

Over 153 thousand census questionnaires were mailed to all households in Malta about two months before census night. About 93 per cent of all questionnaires were collected throughout the official data collection phase, which spanned from 7 November to 4 December 2011. Any remaining questionnaires were collected during the follow-up phase, until a final 96 per cent coverage rate was reached.

Throughout the process, almost 1200 persons, headed by the Census Officer and two deputies, were recruited for data collection purposes. A group of six district managers were responsible for ensuring the smooth running of the enumeration process within each district. In addition, five area supervisors were responsible for the transportation of all census questionnaires from 41 schools across Malta and Gozo which provided logistical support to enumerators during the census period to the premises housing the census office prior to the keying-in of data.

Each enumerator was required to completely canvass the assigned area and record all persons, together with private and vacant dwellings in the allotted EA. Enumerators were requested to record all tenements in their assigned area, by listing them systematically in a number of administrative forms which were provided. Each tenement was assigned a three digit serial number which, together with the locality MGC code and EA number, formed a unique identifier for each dwelling. Information about institutional households was collected by the Census Office.

Methodology

The process entailed the identification of street centrelines as linked to an electoral register that had been geocoded employing a street-centerline methodology. Each previous Census enumeration area was reviewed for its population consistency and an adherence to the maximum population allowed in each area. Issues identified in this process were related to the process that instead of identifying a homogeneous area. An enumeration area that had less population was adjoined by another area, which sometimes was at a distance from the main area. This issue introduced some error and in cases eliminated the possibility for inter-Census analysis. However, for the purposes of data capture and population take, the entire population was served, As a potential future Census runs, a street centreline or a point-based analysis could eliminate this issue.

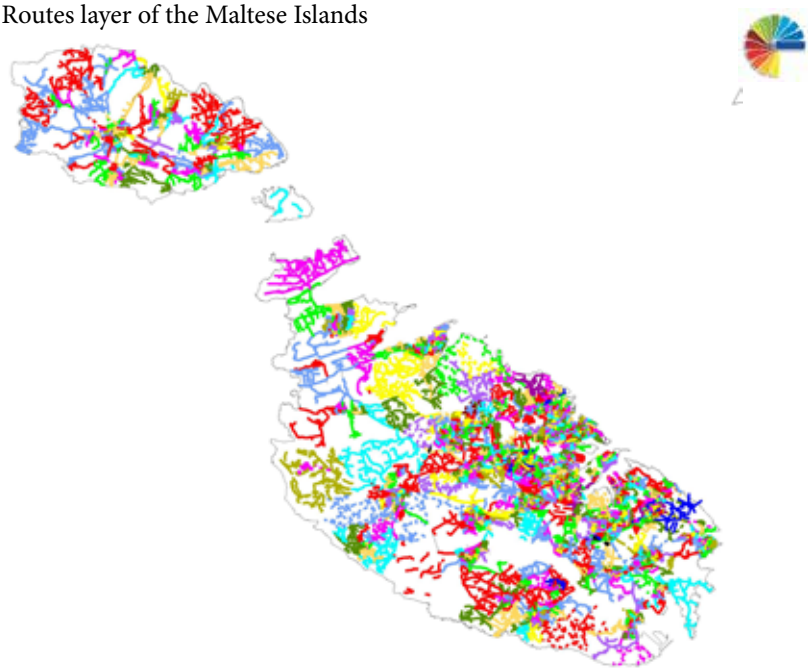
Polygonisation of the Enumeration Areas

The spatial element of the census was brought in through the polygonisation of the Enumeration Areas (EA) an exercise which was carried out for the census in Malta. An enumeration area code represents a series number of streets or routes which make part of the whole council. The Maltese Islands are divided in EA as it is obligatory that every household complete the questionnaire while it is also made feasible for the enumerator to complete all the EA in her/her responsibility in a reasonable time-frame. A registered enumerator is given an EA so that s/he can help the household owner or family to complete the questionnaire in the correct way, if help is needed. Each council was divided into several

numbers of EAs, as this is directly affected with the number of habitable households in the streets. Such distinction can be clearly noticed with the North and North-west areas of Malta with the Southern, Centre and Harbour region of Malta. This was deemed an important step towards a more spatial view on the resulting statistical outcomes. The process incorporated in obtaining such result is presented through the application of the Geographical Information Systems (GIS) knowledge, with the use of MapInfo program.

The routes of the census as used by the enumerator were uploaded and a thematic map for each council was created in order to visualise much better the different routes. In the thematic map is was necessary to have each route number in the council in different colour in order to be easily distinguishable from each other. The preparation of the route map, makes it more feasible and detailed enough to have a systematic approach while plotting the polygon around each EA. The delineation of the polygon needs to be accurate as there could be more than one straight line which needs to be grouped together. Another step before starting the polygonisation was to create a new MapInfo layer, precisely in which the delineation of the EA was plotted.

Figure 1 - Routes layer of the Maltese Islands



Source: Planning Authority

The attribute table of the latter layer included three fields in integer type; Council, Eas_Code and CounEas. These became useful in the process of polygonisation and identification. Finally, once the above preparations were completed, the user must load all the three layers together in one session and save them as a workspace. This will help the plotter to have all the layers aggregated together while it saves where the plotter would have stopped before closing, since this exercise requires a certain amount of time to complete all the coverage of the Maltese Islands. In the workspace, the order of the layers were as follows; routes layer, the thematic map for the polygonisation, the council layer and the orthophotos. The council layer, was necessary for reference of the local council boundaries.

For a more systematic approach to the polygonisation it was better to start from the Northern area moving towards the South of Malta. This helped in the visualisation when delineating the polygon. Once the polygonisation is initiated, it was important to take note of the routes that have the same number, in order to group them together. One must also take into consideration the dynamics of the area, this was aided through the addition of the orthophotos in the workspace. This helps to indicate things such as the width of the road and the building environment amongst other features. Including such layer keeps the perspective of the EA in a more realistic scenario.

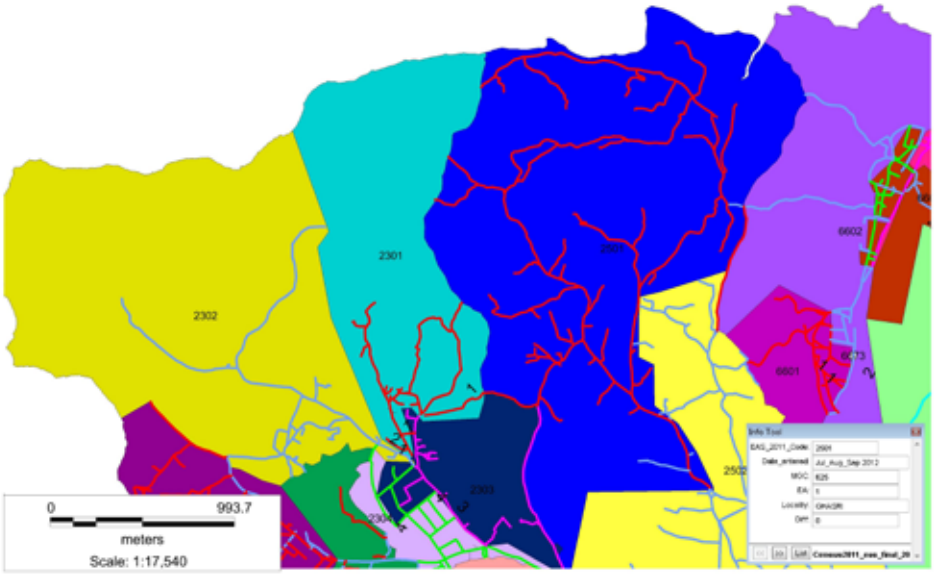
When creating the polygon it was very important to be vigilant with the exact roads that indicated the number of the EA upon the roads themselves. The number upon the roads was made visible through the label feature found adjacent to street layer in the control layer. This was an aid in the process so that one can continue delineating the polygon with more ease and technological facility. While doing this exercise, it was also crucial to see where each point was being plotted, as the polygon was plotted matching with both the outline of the base map and the council layer.

In completion of the polygon, it was important to fill in the information required for each polygon so that each polygon has a unique identifier, whilst no information will be lost. This identification is important to be done while carrying out this process as it will appear in the information tool. Every route polygon includes two components. These are the EA and the council code. After creating the polygon these two identification numbers combine in order to produce the enumeration area system 2011 code. Every enumeration area system 2011 code is unique for every polygon created. Taking as an example, Table 1 shows local council Attard, in the attribute table of the polygonisation layer, one need to input the information accordingly.

Table 1 - An example of the information to be inputted for the polygon

Council	EAS_Code	CountEas
12	**01**	**12001**

Figure 2 - Inputting the information of the polygon

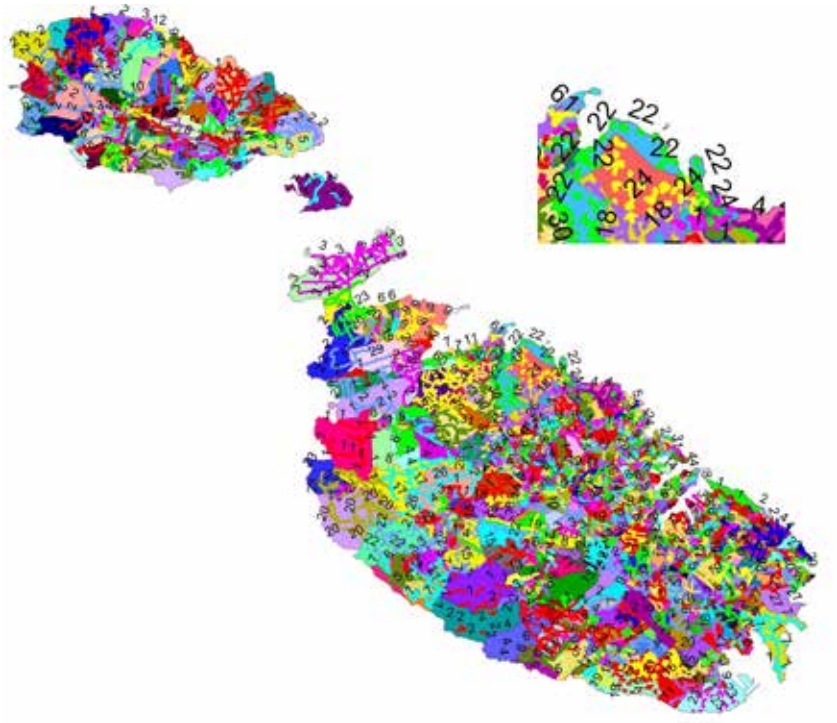


Source: Planning Authority

The information can be obtained from the different reference layers in the workspace. By selecting the information tool from the toolbar of the Mapinfo and clicking on the council, one would be able to see the council number, in this case it is 12. The council number is always constant throughout the council. The EAS_Code, is the route number for the enumerator, each route hold a different number, therefore this will be a unique identifier for the council, and the range of the EAS_Code will vary from one council to another. Finally the CountEas is the Council number and the EAS_Code combined together with a 0 in the middle, this will produce the above mentioned enumeration area system 2011 code.

The process entailed the importance of making sure that proper backups are made regularly, so that if an error occurs during the polygonisation, the plotter would have the latest copy to refer to. An important step in the overall process was error checking. Throughout the process it was important to constantly check the total number of routes from both layers, confirming that the enumeration areas and routes are corresponding to each other. The checking process was repeated for every council. It was important that the main workspace was being updated with every council correctly.

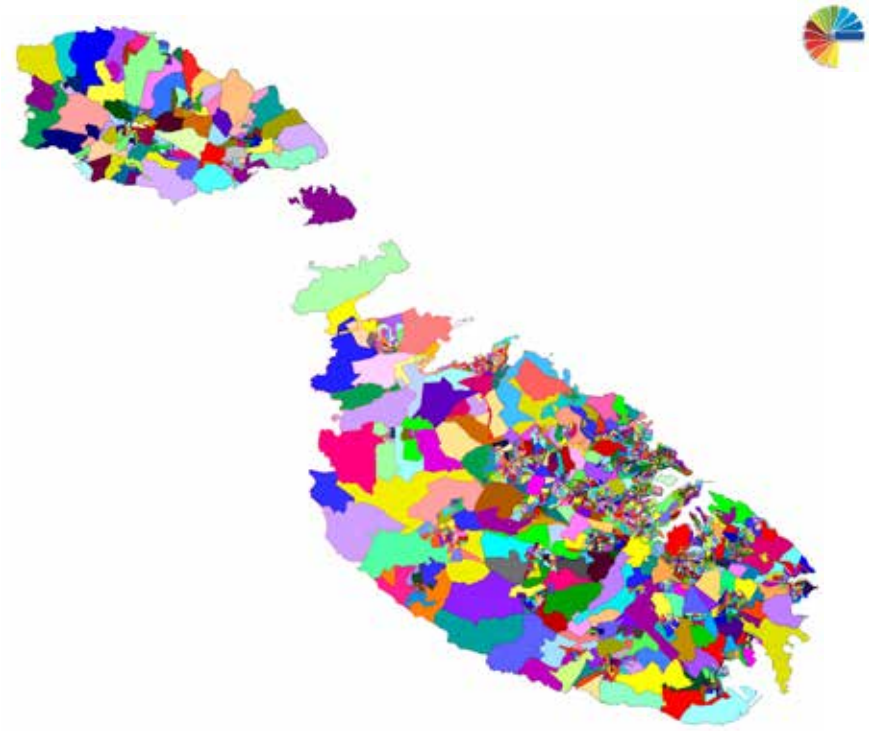
Figure 3 - Delineated EA's Final Map



Source: Planning Authority

The final result was a visual display delineating 1021 polygons representing all the enumeration areas of Malta and Gozo, furthermore the output can be a valuable aid to coordinate more professionally with what happens during the census year. In defining such information the process can be further on utilised by governmental entities to carry out further studies and research.

Figure 4 - Final Census 2011 EA's Map



Source: Planning Authority

Integration of textual data was carried out through a join function with SQL that enabled the tabular data extracted from the Census database, which data (flat table) was amalgamated with the spatial data, in turn allowing for both dataset integrity analysis and eventual analysis across the thematic areas (Formosa, Census 95, 2002).

Data on population and housing could be thus analysed in terms of both descriptive statistics and also spatial statistics, where the data was given an added-value where locational data was generated into choropleth maps, graduated maps and should the point data be available, heat maps (Formosa, Malta National Statistics Office Maps, 2014). Each of these categories was related to the map.

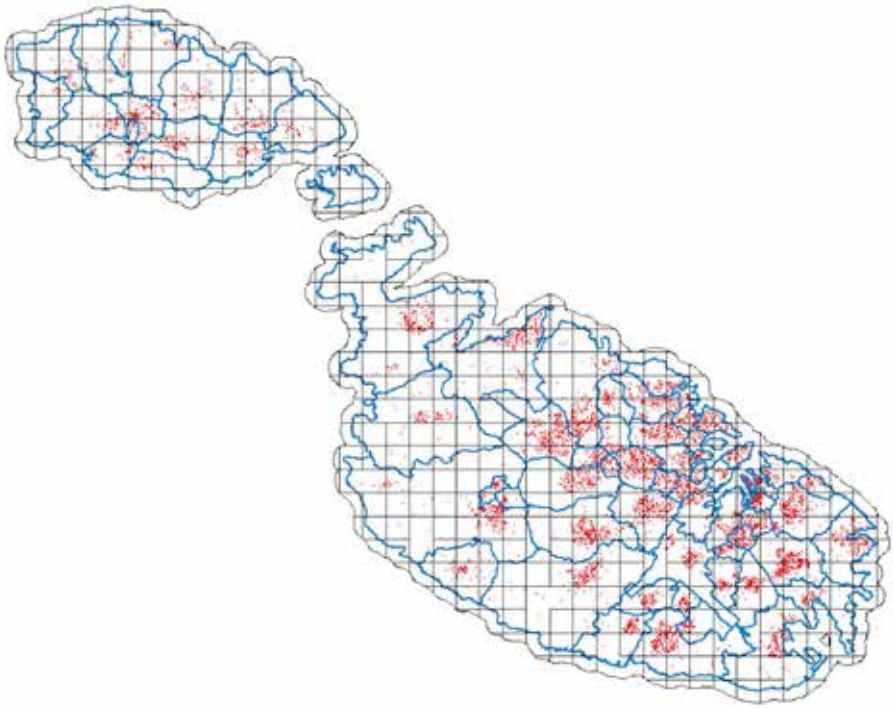
Variables for Analysis

The next step for Census data analysis pertains to the creation of various data structures that would enable the analysis through various methods: mainly the current polygon based structure, a grid-based structure and a point-based structure.

Whilst the polygon-based structure has its advantages in terms of being organic such that the polygons show change over time in terms of population change and movements, it has its limitations in that each Census take requires either a full redraw or major amendments and editing due to change on the ground (Formosa, Censiment 2005, 2005). The grid-based structure as debated at EU-level strives to create conformity across the countries in that it is based on a 1km squared base that would not change over time as it is based on a geographical construct as against a population construct (Figure 5).

This each cell would house a number of streets or households. Should data be only available at street level, then one should be aware that a number of cells could have no population, whilst others could acquire more than their real share. This is due to the context of centroidal function, where each street is designated a central point called a centroid and all population in that street is located in that areal point.

Figure: 1Km Grid base layer



The other method would be based on a point-based spatial layer where each dwelling unit would be located in real space and where the grid-analysis would be easier to conclude and as each unit is located in its point in space, then aggregation of all points in that street in a centroidal location is eliminated. This way the analysis would be carried out in a veritable methodological process as the data is correct in terms of where people live and where buildings are located. In addition to grid map analysis, heat maps could also be created for each parameter (attribute) that is designated to that point and analysis can be built up for any layer type and aggregate. This method spans boundaries as designated by local councils or districting as each point can aggregate to an adjacent point should they be of mutual structure and composition. A collection of points could span a number of streets or localities, thus creating new areas of analytical interest. The Census methodological process is taking a new step forward in that it is analysing new forms of thematic activity.

The next stepped approach to GI and Census

The NSO applied and was awarded a project entitled "GEODE: DataCycle process for the Spatialisation of Fundamental Spatial Elements for Census Implementation ", which aims to achieve an undertaning of the realities pertaining to the creation of a fundumental dataset based on address point data that would be geocoded and offer a basis for the creation of more aggregated data that would allow for inter-thematic analysis ranging from Census through to national and international data analysis.

The main aim is to study, understand, mindmap and map the existing and proposed geocoding processes employed and how they are currently employed and in turn could be used for future data integration based on Census and relative dataset creation. The study will take up the recording of the different data structures that exist in non-spatial and spatial formats inclusive of point, line and area vector data or rasterised data structures.

The project will identify the main datasets, mainly from Census 2011 and other key spatial datasets that could be identified as case studies for the mapping of a point address point dataset, as against the current random creation of boundaries pertaining to unique and at times insular pourposes, such as the creation of a new administratice council boundary. The project would study how far one can go in order to create the base addresspoint dataset, review the current datasets, identify methods to integrate and where non-existent, create new datasets. The process will investigate how the datasets could be integrated, how that integration would aid the data transposition process, map the addresspoint datasets, enable the creation of higher-level polygonised data layers and in turn identify the best technologies the creation, conversion and retention of such a base dataset and point towards the creation of polygonised or raster-based models.

The project structure ensures an investigation into the spatial data availability and

implementation processes of the systems currently in place and new ones created or integrated within the project. The project is expected to deliver an implementation-oriented document based on the datacyle lineage structure that will focus on the DIKA model (Data-Information-Knowledge-Action) that would enable the NSO to create the basis for its geocoding Census knowledge and human capacity as well as empower persons to create, use, interpret and publish. The results will create a series of procedural, methodological, visual and analytical data aids for the entities creating the geocoded layers and which will partake to the inter-thematic analytical approaches.

The results will create a knowledge-based guidance document highlighting the best datasets to employ in such a system through the delivery of: a) the creation of a geocoding process and structure for the base Census data. Another result entails the creation of a GIS-based system in the NSO and training thereof; and b) a series of process lineages, implementation manuals and training sessions. The results will enable the NSO to build a spatial-based capacity and the general public to access data in a visual dimension.

Conclusion

The NSO is Malta's national authority responsible to carry out the census every 10 years, the process requires certain amount of time to be prepared, disseminated, collected and then manually inputted to carry out the statistics as a result of the collected questionnaires. This national exercise deliver huge amount of valuable data which until now was only limited to the hardbound books created as a final product, once the statistics are finalised. Technological advancement showed that a far more valuable use can be made from the data acquired in the census. GIS was the answer to introduce census data in a spatial platform.

GIS has clearly demonstrated its relevance in such process producing results which were generated from the NSO data. The stages taken into consideration along the polygonisation was a chain reaction of all the other datasets. The preliminary review of the data needed and availability reduced the time taken to arrange the datasets as they were already created. As such, the methodology employed in the exercise served as an aid for the whole process to be plotted clearly while making sure to take into consideration the boundaries of the other reference datasets. The resulting map visualised the polygonisation of the EAs. The utilisation of the output later is not limited only to the GIS users but also experts in other fields, serving them as an aid for a more local representative perspective. The importance is referred to the fact that incorporating their data to the EAs layer, will inherently have a higher level of detail than the regular local council boundary.

References

Formosa, S. (2002). *Census 95*. Retrieved from Web-mapping the 1995 Census of the Maltese Islands: http://www.mepa.org.mt/Census/index.htm

Formosa, S. (2005). *Censiment 2005*. Retrieved from Censiment 2005: http://www.tcnseurope.org/census/2005/

Formosa, S. (2014). *Malta National Statistics Office Maps*. Retrieved from Malta National Statistics Office Maps - District and Locality Thematic Data: https://mangomap.com/nsomaps/maps/23649/malta-national-statistics-office-maps#

National Statistics Office. (2012). *Census of Population and Housing 2011: Preliminary Report. Malta*: National Statistics Office.

National Statistics Office;. (2014). *Census of Population and Housing 2011: Final Report Malta.* Malta: National Statistics Office.

United Nations Economic Commission for Europe;. (2006). *Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing*. New York and Geneva: United Nations.