

POPULATION AND HOUSING CENSUS MALAYSIA 2010: DEVELOPMENT IN TECHNOLOGY AND APPROACH

Rozita Talha, Prema Letha Nair and Muhammad Fadzil Ismail¹

Abstract

Development in technology has opened a multitude of opportunities for research across a wide variety of disciplines. The ever increasing demand on reliable statistical data has been accentuated by information revolution, new approaches, methodologies and other challenges. This paper will undertake to discuss the various areas where developments in technology and approach that will be used to enhance the implementation of the census. The Population and Housing Census is a total coverage and complete count of the population, household and living quarters at a given time. The paper highlights the enhancement in the utilisation of Geographical Information System (GIS) in the mapping process, multi-modal approach enumeration, the data processing activity and data dissemination. These are among the initiatives and strategies that will be deliberated in the paper which will further strengthen the efficiency and effectiveness of the census taking.

Keywords: Geographical Information System (GIS)

Introduction

The Department of Statistics Malaysia (DOSM) has had a long experience in undertaking national population and housing censuses, apart from its responsibility of collecting, interpreting and disseminating statistics in other fields. Censuses had been undertaken since the 19th century for the years 1891, 1901, 1911, 1921, 1931, 1947 and 1957. Since the introduction of the Census Act 1960, DOSM has conducted the Population and Housing Censuses of Malaysia for the years 1970, 1980, 1991 and 2000. The next Population and Housing Census of Malaysia will be conducted in July 2010. Generally, the task of undertaking the Census is done once in every ten years.

The Census² is the most important statistical collection undertaken by the Department of Statistics Malaysia in terms of the coverage and massive operation. The prime role of a census is a total coverage and to accurately count and measure the numbers and key characteristics of the population, households and living quarters in Malaysia at a given time that is on Census Day. The Census provides a reliable basis for the estimation and projection of population, frame for other surveys, distribution of government funds and other related use that requires population as a base. This undertaking represents the principal means of data collection aimed at providing a comprehensive set of statistical information for planning, formulation and

¹ Rozita Talha is currently the Director, Prema Letha Nair is currently the Deputy Director and Muhammad Fadzil bin Ismail is currently the Assistant Director of Population and Housing Census Division, Department of Statistics.

² From this page onward, Census refers to Population and Housing Census.

implementation of government policies and programmes. The key strength of census compared with many other statistical collections is the ability to provide population data at micro level.

In line with the advancement in technology, the next census will see the advent and importance of technology in the approach and methodology that will be used in undertaking the census. Looking at past experiences and benchmarking best practices, new approaches and technologies will be adopted to further enhance and strengthen the undertaking of the census.

This paper will discuss on the following:

- Enhancement of the utilisation of GIS in the mapping process;
- Multi-modal approach where for the first time a self-enumeration and internet based forms will be used;
- Adoption of Intelligent Character Recognition (ICR) technology to process data;
- Computer Assisted Coding (CAC) where electronic data dictionary will be used instead of flipping through coding manuals;
- Use of Short Message Service (SMS) for Census publicity and operation;
- Dissemination of data through web based portal; and
- To develop a central repository database for population.

Utilisation of GIS in the Mapping Process

The advent of computers as information processing tools and the development of GIS have measurably assisted the geography discipline. GIS was initially implemented in the government agencies about two decades ago notably in surveying activities. Today, there is an industry being created around maps and specialised portals being launched to enable the use of geospatial information in day to day development activities. Disaggregated data or small area data or micro-data, as well as digital maps meet the need of a host of new data users.

GIS involves the digitising of census geostatistical administrative areas. It enables the integration of analysing geographic spatial data with mapping software where it captures, stores, analyses, manages and presents data that is linked to location. Basically, GIS is considered as an application tool in terms of spatial analysis and distribution.

The application of GIS is often classified into three levels (Arnold & Applebaum, 1992):

- Use GIS to do simple things that we have always done;
- Use GIS to do complex things we seldom or never do; and
- Use GIS to do brand new things that revolutionise or thinking and creating new hypothesis.

GIS can be seen as a software package for the storage, analysis and presentation of geographical information. In recent years, GIS has collected impressive

technological partners to its repertoire, such as satellite driven remote sensing system and global positioning systems, computer assisted cartographic systems, computer assisted drawing systems, and of course global internet infrastructure.

The geographic information process consists of three stages; data acquisition, data processing and data dissemination. GIS technologies were increasingly used in the field of data collection, including census mapping process.

Census processes entail three stages; namely pre-census, census and post-census. Mapping, as an activity of pre-census, has been an integral part of census-taking for a long time. In the mapping process, the whole country is subdivided into enumeration blocks³ (EBs) to ensure full coverage and facilitate census operation. The various geo-statistical levels can be subdivided into two main types;

- Administrative areas – gazetted areas used in the census, such as administrative district (AD), mukim, sub district and local authority area (LAA); and
- Statistical areas – defined by DOSM for census operation such as EB, census circle⁴ (CC) and census district⁵ (CD).

Maps are important tools in data collecting and monitoring census activities. CD maps are used by Assistant Commissioner and District Superintendent to plan and control census operation on the field. Maps allow the census officers to strategically plan, make assignments, identify problem areas and implement remedial actions quick. EB maps are used by enumerators to carry out their fieldwork and as a guide to the area assigned to them. Each EB map is accompanied by boundary description to help enumerators to easily identify the area assigned to them. For the purpose of operation control, seven EBs will form one census circle (CC) and will be supervised by a supervisor. The maps are essential in alleviating the left out or overlapping areas especially in the rural areas where the land marks are not easily identifiable.

The history of GIS development in DOSM had started with its first baby steps in 1989⁶. Issues pertaining to the dissemination of census data in various forms and the availability of new technologies (hardware and software) were studied. Subsequently, a GIS consultant from the US Bureau of Census was engaged to assist the Department in developing and implementing this technology. The process of digitising all of the 40,465 EB maps was started in 1994 and completed in 1996. During 1997 – 1998, GIS Unit completely finished the digitising process for EB boundaries and data attribute⁷ using the 1991 census data. In 2002, DOSM continued to make adjustment and updates on 1991 EB boundaries to Census 2000's boundaries which involved 52,877 EB's and 300 fields of data attribute. The task was completed in 2004.

³ EB is Each EB map is restricted to the preferred range of 80 – 120 living quarters (LQ) or 500 – 600 people.

⁴ CC is ...Every CC consist of 7 EB's.

⁵ CD is Every CD formed will consist of 100 – 120 enumeration blocks (EB).

⁶ The development of GIS is based on a paper written by Mr Kho Siew Lon.

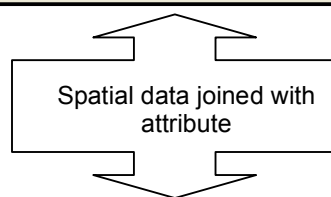
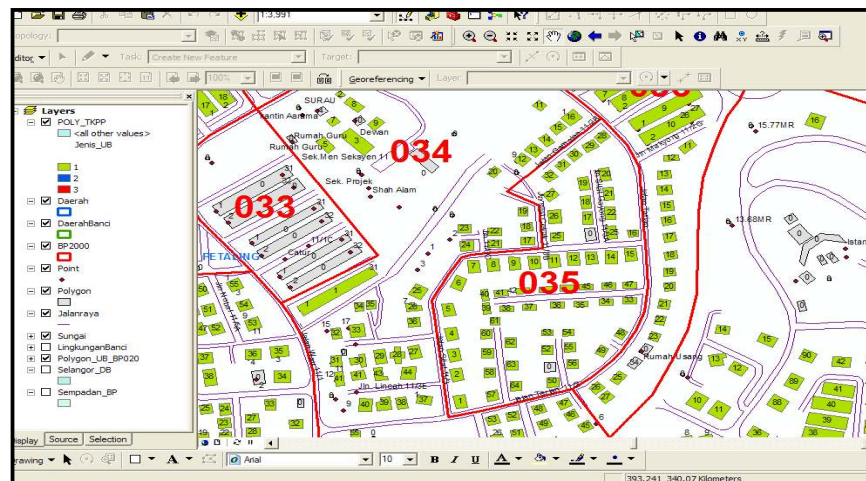
⁷ Attribute – database.

As a result, the Department has successfully produced Census Atlas Report 1991 and 2000 as well as small area statistics for purpose of analysis as required by stakeholder and users.

Prior to 2006, GIS development in DOSM was concentrated only at the head office. In planning for the 2010 census, DOSM expanded the use of GIS to the state level. State offices were equipped with GIS facilities and extensive trainings were provided for the personnel at the state level.

The enhancement of GIS will benefit DOS to produce EB digital maps. User of EB digital maps will be able to find, manage, retrieve and restore the EB in a much more easy manner as all the data is stored in the GIS database. Furthermore, the implementation of GIS will also benefit the management by the Methodology Division. Monitoring and controlling of all the updated EB digital maps and attribute data becomes more efficient using GIS.

Figure 1: Census database (joining spatial data and attribute)



NO	UB	TX	NAMA	NO	TINGKAT	BONKONG	LOKUS	KOTAM	BANDAR
1043005	003	2155	TULUS AMAN SDN BHD	3A &	TKT 1 & 2		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	003	2155	TULUS AMAN SDN BHD	3A &	TKT 1 & 2		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2290	TUAN TANHER SDN BHD	05-01	TKT 5	PLAZA MASALAM	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2290	TUAN TANHER SDN BHD	05-01	TKT 5	PLAZA MASALAM	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2201	TY TELECOMMUNICATION	LOT-H	TKT 1	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2201	TY TELECOMMUNICATION	LOT-H	TKT 1	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2236	TRY 'N' TAKE COLLECTION	LOT-L	TKT 2	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2236	TRY 'N' TAKE COLLECTION	LOT-L	TKT 2	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	002	2048	TRUST LIGHTING COLLECTIO	33 &	TKT 0,1,2,5	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	002	2048	TRUST LIGHTING COLLECTIO	33 &	TKT 0,1,2,5	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	002	2072	TITAN UMMI ENTERPRISE	178	TKT 2		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	002	2072	TITAN UMMI ENTERPRISE	178	TKT 2		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2245	TME GALARE	LOT-L	TKT 2	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2245	TME GALARE	LOT-L	TKT 2	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2219	THE ONE MUSIC HOUSE	LOT-L	TKT 2	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2219	THE ONE MUSIC HOUSE	LOT-L	TKT 2	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	004	2166	THE BAKER COTTAGE	13	TKT 0		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	004	2166	THE BAKER COTTAGE	13	TKT 0		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043003	001	2001	TERACHI DESIG	36-1-	BLK 36	KEMUNCAK SHAH ALAM KONDOMINI	JLN 9/2	SEK 20	SHAH ALA
1043003	001	2001	TERACHI DESIG	36-1-	BLK 36	KEMUNCAK SHAH ALAM KONDOMINI	JLN 9/2	SEK 20	SHAH ALA
1043005	005	2278	TEN & TEN FURNITURE CENT	LOT-L	TKT 4	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	005	2278	TEN & TEN FURNITURE CENT	LOT-L	TKT 4	SHAH ALAM MALL	JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	003	2156	TARGET PHARMACY (SHAH AL	5	TKT 0		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	003	2156	TARGET PHARMACY (SHAH AL	5	TKT 0		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	004	2168	TAGYUDON & PARTNERS	13A	TKT 1		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	004	2168	TAGYUDON & PARTNERS	13A	TKT 1		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA
1043005	002	2181	TAKAPU BILAS (U)	3A	TKT 1		JLN TENGGU AMPUAN Z	SEKSYEN 9	SHAH ALA

Figure 2: GIS Census Database

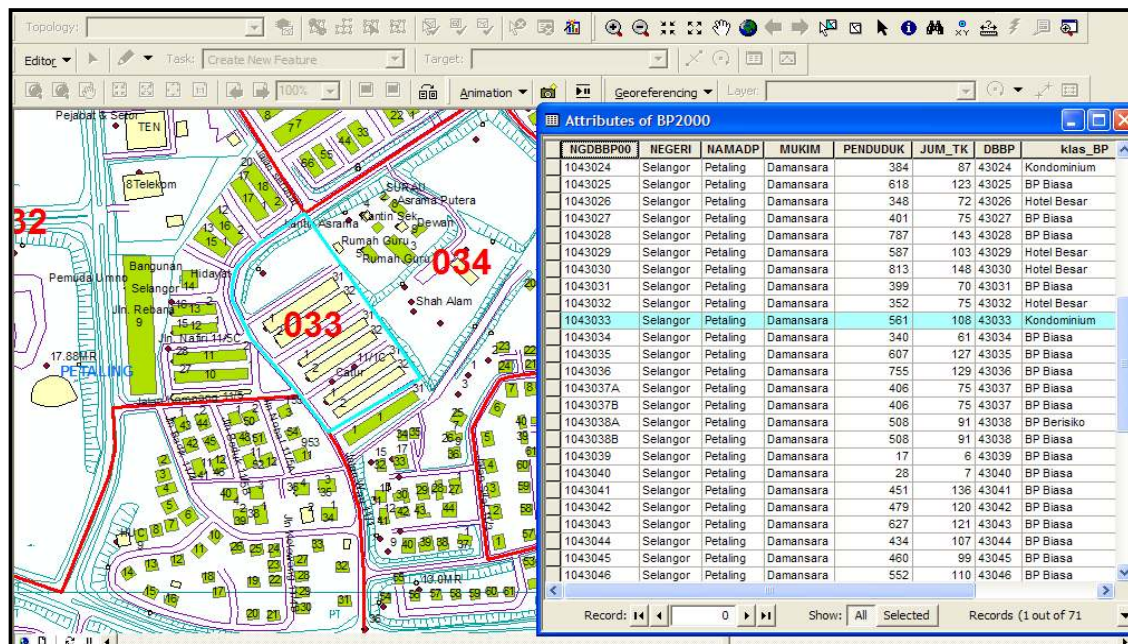
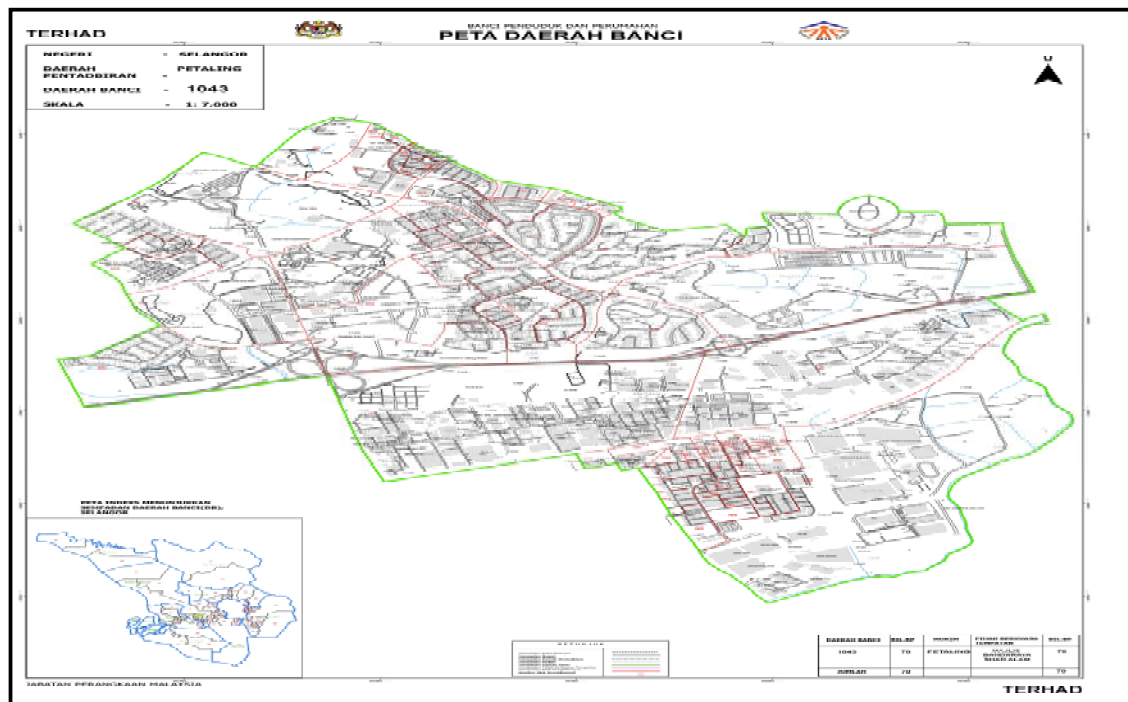
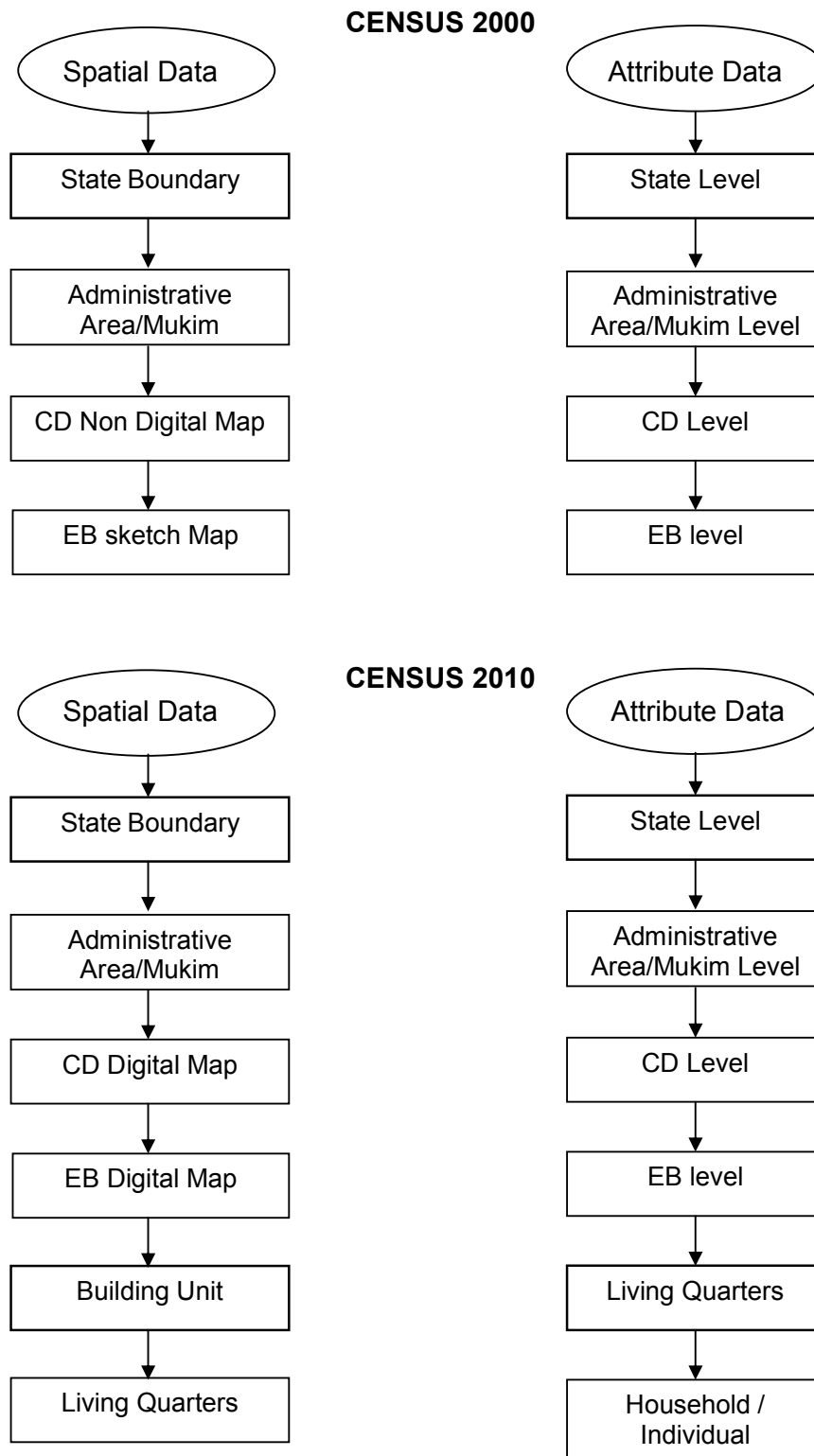


Figure 3: CD Digital Map



The following Chart 1 shows the flow for spatial and attributes data in Census 2000 and Census 2010. Comparatively the GIS application as planned in Census 2010 will assist in creating a small area statistics up to living quarters (LQ) and individual. In other words all the LQ will be represented for both spatial and attribute data.

Chart 1: Spatial and attribute data flow for Census 2000 and 2010



EB digital map will be used for the 2010 census. Updating of EB maps using GIS software was fully implemented in 2007.

Figure 4: EB Sketch Map for Census 2000

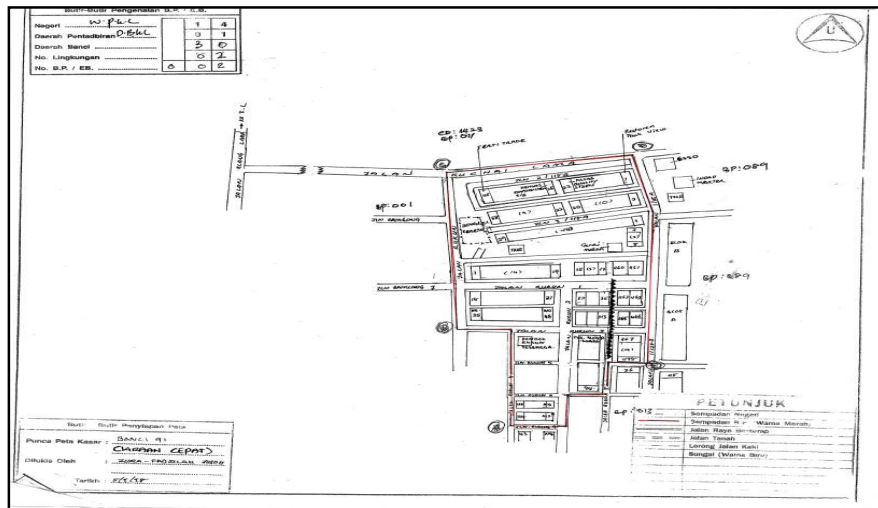
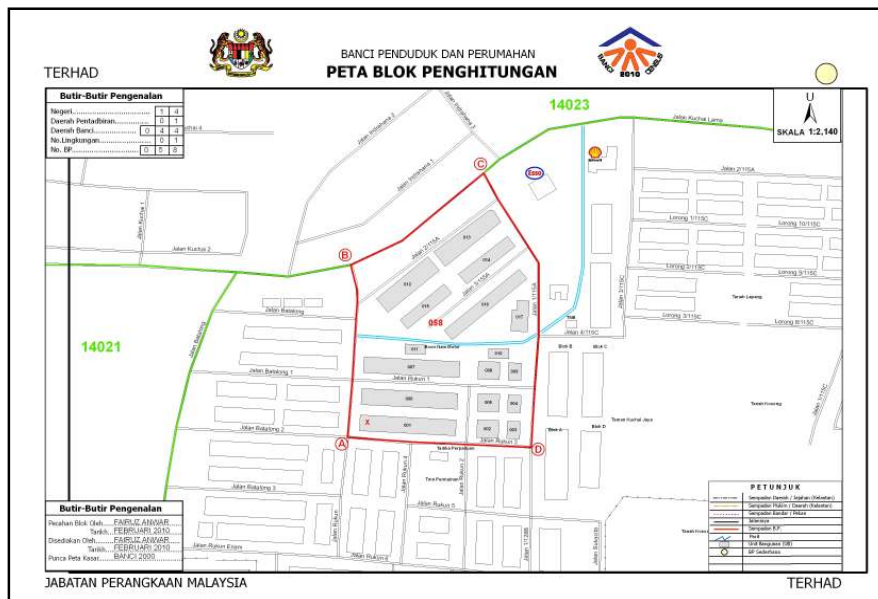


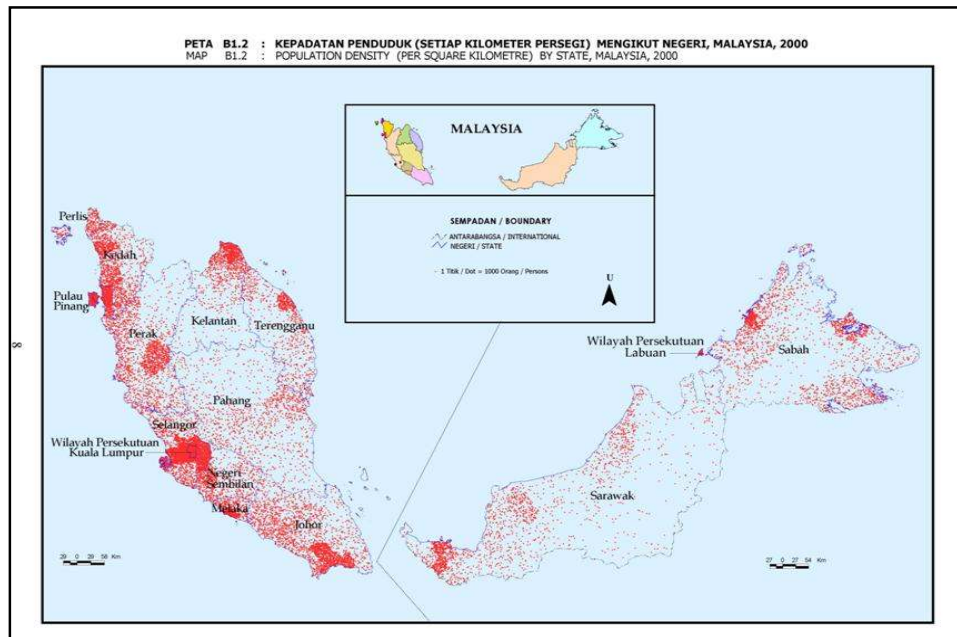
Figure 5: Digital EB produce using GIS Software for Census 2010



The GIS software is used to process and construct the different levels of census geography. Small area statistics from census data which consist of demographic data, living quarters, household, facilities and others can be derived through GIS application. Some of the products generated from GIS application are highlighted below.

Population/demographic aggregated database derived through buffering process and aggregated based on the number of EB is capable to generate data by radius as shown in Figure 3 and 4 as well as the overall population density of the whole country (shown in Figure 5).

Figure 8: Population Density per Square Kilometers, Malaysia, 2000



Moving forward, the DOSM will delve into more innovative usage of GIS for data generation, presentation and dissemination. Exploration in these areas would bring the DOSM into the next level of data production in terms of speed, type of data and detailed information such as 3D presentation, GIS small area modelling, and estimation.

Evolution of ICT utilisation in Census 2010 Data Processing

DOSM is one of the pioneer agencies in the use of technology among the government agencies in their daily work. The Optical Mark Recognition (OMR) technology was used during the 1970 and 1980 censuses. For the 1991 and 2000 Censuses, DOSM opted for manual data entry system in order to enhance the usage of personal computers among the personnel as well as a change in the strategy towards decentralised processing.

Census processing has always been a challenge in complying with the standards and requirements vis-à-vis the volume of questionnaires that have to be processed. In order to fully leverage the use of ICT in data processing, ICR technology will be introduced in the next census.

The census forms will be scanned using ICR machines at the seven processing centres throughout the country. Specialized computer software will be used to interpret the handwriting on the respective images of the census form and transforms them into a single computer data file. It is anticipated to reduce processing time and hence reduce the number of staff needed when compared to the huge number of data entry clerks used in the previous censuses.

Manual coding of occupation and industries usually requires time and effort for finding the right codes. Computer Assisted Coding (CAC) provides friendlier and easier search function. Pre CAC edit specification allows the basic checks be done electronically. Once the forms passed pre CAC edits, they will then go through further onscreen and online coding processes for education and migration.

ICR technology entails onscreen quality control checks so as to ensure the accuracy and reliability of the data. To further enhance the validity of the data, quality checks will be performed on all members of the household in the respective EBs if more than 50 per cent of the industry or occupation codes are found to be inaccurate. Onscreen quality checks increase the speed and efficiency of the processes.

Online web based tabulation

With the innovation of ICT and new software packages, the processing activities can be monitored not only at the processing centres but also at the headquarters. This will enable to alert the census operation centre in headquarters as well as the processing centres at state level of the bottlenecks. It also allows direct monitoring of the census progress by the management or the commissioner office. The processed data at the processing centres will be transmitted online to head office for tabulation processes. Data mining facilities in place of the traditional mainframe processing will be used and this will expedite the production of census output and releases of Census products.

e-Census

Completing the census questionnaire over the internet is used by many countries around the world. The e-Census approach was found to be effective to capture respondents mainly in urban areas and gated communities. DOSM for the first time will introduce this approach in the 2010 census with the aim of leveraging on the internet savvy community. With the increase in the accessibility and the capacity of broadband in the country, at this juncture, it is the right time for the Department to capitalise on e-Census.

Strict security measures at various stages of process will be imposed to ensure confidentiality and safety of information provided by the respondents. The passwords and pin numbers will be given to households during the visits by enumerators. Forms completed via e-Census will be processed electronically once received via the Internet. The notification of e-Census submission will be automatically sent to the field supervisors via Short Message Service (SMS) for field collection monitoring process.

- Built-in checks will be installed to monitor the quality of data.
- Status tracking (monitoring) of the form will enable the field operation centre to provide a reminder to the respondents as to the need to complete the questionnaire within the specified time. If this is not met, steps will be taken to visit the household to conduct an interview.

Census metadata, census data and GIS will be integrated in the *National Enterprise-Wide Statistical System* (NEWSS) which is being implemented under the Department's ICT Strategic Plan (ISP). NEWSS will provide the platform for an integrated statistical system for collection, processing and dissemination of statistics. This platform will be used to facilitate future census undertaking. Census data for the years 1991, 2000 and 2010 will be kept under NEWSS repository which will harness future data request needs.

New Approaches in Census 2010

Drop-off and pick-up (DOPU)

In view of the changing social and demographic pattern of the population, for the 2010 census, DOSM will expand the approach for census taking by introducing self-enumeration. Self-enumeration using hard copy and electronic questionnaires will help to improve response rate and is targeted at urban areas. Face to face interview will be used in both urban and rural areas. The duration of the census enumeration will be expanded from two weeks to six weeks. This multi-modal approach will ensure the completeness of the census coverage.

Population and Housing Census is a national project. Enhancement of collaboration with other Government Agencies (OGAs) in conducting the census is important to ensure the smooth running of the project. For the 2010 Census, DOSM will fully utilize other government agencies' facilities in terms of transportation and other logistic facilities as well as man power in the process of data collection.

Census publicity

The objective of census publicity is to provide information, awareness and understanding to the public on the purpose and importance of census taking.

Census publicity can be very costly. However, a wide national coverage is important to ensure public awareness and to gain cooperation from the general public in providing census information. Among the strategies formulated to ensure effective communication are through publicity campaign, announcements, advertisements, minister's launching, full media coverage, press conference, census brochures and pamphlets as well as publicity through electronic media such as internet and mySMS.

A census portal has been developed which contained general information about the conduct of census and later will also include aggregates of population database. The portal can be accessed through the departments' website at www.statistics.gov.my.

Conclusion

ICT has become an essential element of the resource management and data dissemination. Rapid development in technology has changed the scope of information processing whether as end user application or for technology support. In addition, the advent of efficient computing technique, based in advances in hardware technology has provided a boost to the planning, time, decision making process and to achieve the customer/user expectation.

For Census 2010, DOSM has enhanced the usage of ICT in the mapping process using GIS, data processing via ICR and also data collection via e-Census. As for publicity, a census portal has been developed which contained general information on census and inter 2010 census data.

To cater for the changes in social and demographic pattern of the population, drop-off and pick-up (DOPU) will be introduced in the coming 2010 Census.

It is hoped that the utilisation of ICT and new approaches during the 2010 census will further enhance the capabilities and integrity of DOSM in carrying out her duties and responsibilities as the sole official provider of statistics in the country.

References

Handbook on Geospatial Infrastructure in Support of Census Activities; United Nation, 2009.

Principles and Recommendations for Population and Housing Censuses, Rev.1, United Nations, 1997.

The Global Geospatial Magazine – GIS Development June 2008 Vol. 12 Issue 6; December 2008 Vol. 12 Issue 12 and January 2010 Vol. 14 Issue 01.

A paper on “*Potential and Possible used of GIS in DOSM for Post Census Activities*” was written and in 1990, the budget was approved for the development of GIS.

Manual on GIS for Planners and Decision Makers, United Nations, 1996.