	CONSULTANCY FOR UNFPA OFFICE FOR THE ENGLISH AND DUTCH SPEAKING CARIBBEAN (JAMAICA)	**SUMMARY OF REPORT**	
		1. Mission ID:	**CAYMAN0801**
		2. Type of mission:	Support to Data Collection

Expert:		**Report date**
3a. Name:	**3b. Post**	2/6/2008
ARI NASCIMENTO SILVA **arinsilvabr@gmail.com**	CONSULTANT	

4a. Commencement date of mission	**4b. Completion date of mission**	**4c. Length of mission**
20/05/2008	25/05/2008	8 days

5a. Country being supported	**5b. Program Phase**
Cayman Islands	First

6a. Project Code	**6b. Project Title**
N/A	IT Resources for the 2010 Population and Housing Census

7. Initial Objectives of Mission

To conduct an assessment and provide advice on IT issues including recommendations on how to improve the IT system presently used by the Economic and Statistics Office (ESO), especially with reference to the upcoming Population and Housing Census. The TOR, in full, is in Appendix I.

8. Mission Context

8.1 The Cabinet of Cayman Islands officially approved the planning of a Population and Housing Census, which will be conducted in October 2010.

8.2 The ESO plans for about 250 EEAs, each containing about 100 households.

8.3 ESO has already decided to use scanners for data entry.

9a. Objectives (agreed upon in briefing) of the mission

9.1 Undertake an assessment of the IT resources currently in place at ESO in the Cayman Islands and also assess the need for an in-house IT-expert within ESO;

9.2 Provide advice on IT-system (hardware, software) for the 2010 Population and Housing Census, including dissemination;

9.3 Advice on the training needs in IT, especially for the 2010 Population and Housing Census.

10. Activities carried out

General

10.1 First there was a briefing session with Mr. Ronnie Andersson, Chief Statistician, where the objectives of the mission were discussed and more final details were defined.

10.2 A meeting with Ms. Nicole Emmanuel-Jones, Census and Survey Supervisor, about the procedures for the 1999 census, where ESO already had used OMR (Optical Mark Reading). There is not much experience left in the office from that census, as far as programming is concerned, for the data entry, editing and tabulation processes, nor much documentation of how things were done. They know that a REDATAM database was created for data dissemination and census storage.

Preparedness & Infrastructure

10.3 One of the recommendations that are in order, and this is the very first thing one has to make when starting a project as big as a Population Census, is to design its schedule, even in a very preliminary basis, with just a few actions and tasks (Mr. Andersson had done his "homework" and already has it in an Excel format). Then, later on, these broader tasks can be subdivided into smaller tasks, going down as much as possible, up to the operational level, trying, at this time, to establish, for each task, their duration, dependency of previous tasks, responsibilities, and costs. At the beginning, an Excel sheet might suffice (using the columns for the timeframe periods), but later, it should be better to use a specific software to handle that, Microsoft Project being an alternative (see reference in the Appendices). The standard version, designed for managers who do not need to collaborate with others to build their projects, costs around US$600, and it is more than enough: there is no need to buy the Professional version, which is used by resource managers who work with others across an enterprise to share schedules, and costs around US$1,000. The advantage of using a software like this is

that all the tasks dependencies are handled automatically, and if one task is delayed (or shortened) for some reason, the ones that depend on it will be delayed (or advanced) as well. And the software provides also some other tools, like establishing milestones (such as the census day) that have to be kept fixed no matter what. Subprojects might belong to this broader project as well, such as the Census Pilot, for example. Appendix IV has a small example of a Gantt chart done by Microsoft Project using Census Planning information.

10.4 ESO's office is of good standards, well settled and organized, space is not at a premium. For the census, it might be a little tight, considering all the extra space required to store questionnaires in the shelves (Mr. Andersson said he is thinking on buying locked rackets, which is good), the new machines, and also the extra personnel (computer operators, verifiers, coders, etc.). Mr. Andersson said they could use the meeting room, if needed, for the overflow. I would recommend that a more in depth study of the space allotment for the census should be done.

10.5 Timeliness for the census project is very good; ESO is planning ahead of time for October 2010, even considering that they want to do a pilot one year before census day, which brings us to October 2009.

10.6 All the computers at ESO are linked to a broad government network, which is working fine. However, at the time of the census processing, when there will be the need to transfer larger volumes of data back and forth, from the scanners to the interpreters, etc, to establish a local network for the computers involved in the production might help to speed things up, although I am not an expert on this area, and consulting with the government computer manager he said he did not expected any problems, that the network would be able to support the added extra workload.

Operational procedures

10.7 One of the issues that were discussed is the availability of a census questionnaire through the internet. ESO's concern is that sometimes people do not open their doors to the interviewer, but might answer the questionnaire if they had an Online Census Option. The intention is clearly to try to increase participation in the census, and (maybe) to improve data quality. There are two experiences that I have heard about, New Zealand and Canada, and documents from both of them were given to ESO personnel. According to the documents, both were successful. From my point of view, there are some details that have to be taken into account before adopting this strategy, and a direct contact with these countries should be pursued. A good cartography is very important, since every household (who wanted to do it) would need an identification number to access the online questionnaire. Security is another very big issue, all the communications between the respondent and the central bureau must be in a cryptographic format. Timing, also, is another key factor; households that said to be using the internet must be contacted again if they did not answered the questionnaire in a certain time. Organization of all that information must be online in a house-by-house database, in order for the controllers to know the responses. The questionnaire programming must be able to stop at the user's command, and start again later on. A complete and very thorough help system should be designed to assist in every question the respondent might have. Personally I am not convinced that the effort to design an application like that is justifiable, maybe the approach that was implemented at 1999 is enough to take care of the non-respondents (leaving the questionnaire with an envelope that could be sealed and given to the enumerator).

Hardware

10.8 The decision to use scanners for the 2010 census is sound. The alternatives would be: a) a manual key entry system; or b) data collection using PDAs (Personal Digital Assistant). In the case of Cayman, both have some problems that overcome their advantages. Even if the first one (manual key entry) is sometimes recommended for small countries, it has several disadvantages, such as the time consumed for data capture, and mainly, the quality of the whole process (the time should be increased if a verification procedure is to be enforced to guarantee a minimum of error rates). Manual data entry is also used when there is a political need to employ people in the country, and the wages are low, which does not seem to be the case here at Cayman. And for the second (use of PDAs), although there are previous experiences in other countries using this technique in population census (the latest ones in America being Colombia and Brazil) and some trials being done at USA, there are conflicting opinions about it, which means that there is a need to do a more in depth study, and of course, it would depend on each country's characteristics. In this case of the PDAs, there are other issues that should be taken into account, such as the software to be used (besides an ad hoc programming, the only software that is known to handle PDAs for a survey is CSPro, and it has some limitations as not handling more than one question per screen, multiple-answers have to be done in a one-screen per answer basis, instead of handling check marks), PDAs do not offer a good navigation system to look at the questionnaire as a whole (paper still is a better source to browse), and to enter long text (alphabetical characters) in a PDAs (like for occupation and industry) is still a little bit clumsy, even considering the alternatives of the on screen virtual keyboard, the text hand writing facility or the regular keyboard in some PDA models. On the other hand, PDAs should be less expensive on the overall (you do not need to print questionnaires, PDA prices are coming down very rapidly, and there are small 7-inch screen notebooks for under US$200), they should be more accurate (no double marks), consistency checks could be done at the interview, and questionnaire flow is handled efficiently, automatically skipping questions that need not to be answered, etc. Lastly, PDAs should be faster, since it does not require any further data entry, and the data editing and consistency phase should

be kept to a minimum.

10.9 The scanning process has to be carefully dimensioned according to the total census timeframe, and considering all the resources involved, such as the machines to be used (scanners and computers), the software (for data capture and data recognition), and personnel (for the data verification process).

10.10 First of all, working with a single scanner is a very risky situation because there would be no back up if it breaks during the data capture. So, no matter the scanner's speed (see next item), ESO would be safer if there were two scanners in the office for the 2010 census, preferable with the same brand and model (same spare parts, manuals, operational procedures, etc).

10.11 The scanner at ESO is a Canon DR-7580 which is used for the Labor Force Surveys. It has a nominal speed of 75 ppm (pages per minute), or a 150 ipm (images per minute) in duplex mode. Considering that you never achieve such "propaganda" speeds, we might aim at most for a 70 to 80% of the nominal advertised figures, which means about 100 to 120 ipms. By using the same 1999 questionnaire (just for preliminary dimensioning), it has 32 pages per household that have to be scanned (even the blank sheets are scanned, otherwise you have to have a preparatory station to take them out, which might provoke more errors if some real person is eliminated from the batch). So, to read the 25,000 expected households (800,000 pages), using the conservative 100 ipm speed, it would give us about 8,000 reading minutes, or about 134 hours, and, on a 5-hour day basis, 27 days, or about 14 days working double shifts of 5 hours each, using a single scanner.

10.12 Of course this is only a preliminary figure, and those numbers can change, but some considerations can be done. By looking at the two previous paragraphs, the ideal situation would be using the existing DR-7580 scanner as a backup and buying two new ones, maybe of a newer model (faster and more accurate). By looking in the internet, prices of this kind of scanner are about US$5,000 to US$6,000, plus the accessories and replacement parts, such as lamps, exchange roller kits, etc. that are needed (newer models might be a little bit more expensive). Besides that, each scanner would need its own dedicated computer and human operators (for one or two shifts).

10.13 Other computers will be needed to handle the recognition part, and computers will also be needed for the verification stations (fields not recognized by the recognition system, double marks, etc.).

Software

10.14 From the software point of view, and just for data capturing, ESO is using TELEFORM for the Labor Force Surveys, and ESO intends to do so for the next census as well. TELEFORM is very simple and easy to use, not that expensive, and might fit their needs. It has to be supported by some outside programming such as for data controlling (number of households and persons by district and EA). It might be necessary to buy some more licenses, depending upon the number of servers that will be used to recognize the data, either standalone versions, or a workgroup concurrent network version, depending on prices and availability.

10.15 There are other software for document scanning, like SIPO (Integrated Optical Processing System, from Active Software) and EFLOW (Unified Content Platform, from Top Image Systems), among others. These two were already proven in census processing, they are much more than just a simple document scanner software, since they have all the necessary steps from the scanner itself up to the modules that control all the paper and files flow, but they should be much more expensive than the TELEFORM software.

10.16 After the data capture, there is a need to code the alphabetic questions, the more important ones are occupation and industry (and maybe others like religion and migration). This can be handled by an assisted coding scheme that could automatically code a percentage of the answers, and give to the operator some alternatives for the ones the system could not place a code. Some countries have been doing that since the 1990 censuses, like Mexico and Brazil, and the last experience on this subject that I know of is on-going in Peru for their 2007 Population Census. CSPro has a very easy example on how to do a coding process based on a file having all the possible coding entries (see more references to this software later). This step has to be taken into account when designing the overall census schedule.

10.17 Focusing the editing step after data capture (and coding), the next one should be data editing and consistency checking. There are some alternatives that can be used, even designing an in house application (which I strongly not recommend), such as Blaise (from the Netherlands Bureau of Statistics, CBS, see reference in the Appendices), and CANCEIS (from Statistics Canada, see reference in the Appendices). Blaise, although being a general purpose software, it is known for its Computer Interviewing module, which is very good. CANCEIS uses a nearest-neighbor imputation methodology based on writing Decision Logic Tables (DLTs). As far as I know, CANCEIS is being used in Canada and also in Brazil (for the 2000 population census and 2006 agricultural census). And the one I recommend, which is CSPro, for being easy to use, free, fast, and worldwide used. Its strongest features are: a) it is an integrated package, having data entry (which you do not need to use), consistency checking and tabulation; b) you do not need to be a programmer to use it; and c) by

being a very used software, it is very stable and bug-free. Besides that, it is easier to find people that know CSPro (or its older version IMPS). In fact, CSPro is a merged combination of IMPS and ISSA, a package for survey processing used in the Demographic and Health Surveys, DHS, projects).

10.18 CSPro programming is based on a data dictionary describing the input file (the meta data), and procedures attached to each of the dictionary's variables. The edit rules are written in an if-then-else type of programming structure, and the imputation can be done in a deterministic way, or using hot deck matrices controlled by variables that handle a strong correlation with the one being imputed. Several reports are provided to revise the whole process. The tabulation modules are also very easy to use: there is a fast module to get simple frequencies from the variables, just marking the variables from the dictionary list, and a more sophisticated one that produces cross tabulations using a drag & drop kind of facility that designs the desired table.

10.19 Other computer languages such as SQL and packages like Microsoft Access have been discussed to be used by ESO. For the census project, however, SQL might be useful in the early stages where the census data still are in TELEFORM format, before being migrated to flat files to be processed in the editing and imputation phase by CSPro, but I do not foresee any usability for Microsoft Access in the census, unless it is needed as an auxiliary tool for matching, merging or doing any other specific tasks with the census files.

10.20 By looking at the 1999 census publication, it can be seen that there are many variables (especially at the household level) that have very few cases of NS (Not Stated), for example, the Telephone Service as in the Table below: there are only 31 NS answers out of as many as 14,907 households (0.2%). I recommend that cases such as those should be imputed in the 2010 census, using any of the imputation strategies available (for example, hot deck using the last household with the same characteristics), provided that the NS percentage is very low. By doing that the tables would be neater and smaller (suppressing the NS column).

NUMBER OF HOUSEHOLDS BY DISTRICT AND TELEPHONE SERVICE

	YES	NO	N.S.	TOTAL
GEORGE TOWN	5924	2523	20	8467
WEST BAY	2277	636	4	2917
BODDEN TOWN	1608	340	5	1953
NORTH SIDE	314	39		353
EAST END	312	90	1	403
CAYMAN BRAC	554	178	1	733
LITTLE CAYMAN	51	30		81
TOTAL	11040	3836	31	14907

74

10.21 For data analysis and tabulation, ESO is using so far SPSS for its surveys, which is good. REDATAM might be used also, instead, or in combination with SPSS. REDATAM was used by ESO for processing the 1999 census, but apparently its usage was discontinued at ESO for some reason, and now nobody at ESO has any knowledge about REDATAM. Maybe ESO should concentrate efforts in spreading SPSS to all the statistical officers, in order for all of them be able to work with SPSS, and then having a common "statistical language". REDATAM might be good also, in the event that microdata were to be set available in the internet, but I think it would be too many different software for too few people in the office (10 professionals, including the Chief Statistician).

10.22 As far as data dissemination is concerned, at least for what I was able to see, ESO has limited its production by regular paper publications and their "counterpart" in the Internet (publications in Acrobat PDF format files). There is a need to improve dramatically ESO's output, because this is what it is all about: people have the right to have access to as much information as possible, protecting, of course, the confidentiality. There is one piece of software, named PC-AXIS, from Statistics Sweden (SCB), which is very handy for managing aggregated data in tabular format, and I am sure it would fit neatly the needs of ESO for data dissemination through the internet. In PC-AXIS format, a table has two parts, the metadata part (table description) and the table contents (actual figures, numbers). Building a table is very easy; data can be imported from several sources, including SPSS, Excel, or even plain and flat text files. Each table can have several dimensions, for example, age group by sex by marital status by education. From that, PC-AXIS can produce different

types of combinations, from a single frequency distribution of a single variable (hiding the others) or doing any combination of them, at all the geographical levels (country, district, EA) or other geographical subdivisions like Education areas, or Electoral districts. Data can also be presented on a graphical form or a map form, everything being done directly by the end user. After the data is stored in PC-AXIS, it can even help to produce the final publication reports for the census (or other surveys).

10.23 Thematic maps are a very good way to present statistical information, and it should be used much more frequently. In the past, statistical publications consisted of a huge volume (sometimes more than one) of tabular data. Nowadays, things had changed, and the general tendency is to publish more analytical texts, with many commented graphics and maps, the tables being relegated to a minimum, in an appendix of the publication, sometimes in a CD that accompanies the publication. In order to be able to present our data in a map format, the only thing that is really needed is the boundaries (polygons) of each geographical area (districts, EAs, whatever) in a digital file that could be used by almost any software package, even the ones mentioned here (CSPro, REDATAM, SPSS, PC-AXIS). The most common format being used now, which is becoming the standard in the map market, is a format named SHAPE (.shp extension) which was created by the Environmental System Research Institute (ESRI), a leading company on the market. Every institution who handles maps is able to produce its files in "shape" format. So, it is a matter of obtaining the district boundaries (and EAs boundaries, and any other country subdivision that is needed to present statistical results) in a shape format. From then on, tying those files to the statistical census files is not a problem, provided we are using the same identification scheme for the polygons (in the maps) and the records in the statistical files. There are specific software packages to handle maps, the most known and used comes from the ESRI family, named ArcView (US$1,500 Single User License), but I do not think it would become necessary to use such a package, just for creating thematic maps, unless there is a need to calculate distances and really work using topological and GIS information, which I cannot foresee in the near future for ESO.

10.24 Microsoft Visio is handy when there is a need to design system flowcharts and project workflows. An example of what it can do can be found in Appendices V and VI, having, respectively, the workflow of the census questionnaires at ESO, and the programming representation of the Editing & Imputation phase.

Manpower

10.25 A census project is not more complicated than any other project, but it has some specific difficulties and a very large list of different tasks to face and plan ahead, requiring a very tight organization. There are several things that need to be done, starting with the project schedule, space dimensioning, shelving, documents workflow, subject matter discussions (questionnaire contents), people to be hired (enumerators, coders), advertisement, etc. The Chief Statistician is also the Census Manager, but due to his other duties he cannot work full time with the census. A project like that is not a part-time job, it has to have total dedication from the people involved in its management, being totally focused on the project. For this reason, I would recommend that somebody from the office should be appointed as "Assistant Census Manager", to be 100% dedicated to census matters: if there is not such a person in the office, one should be hired. One of the biggest advantages of this approach is that this person would be the "focal point" for all the census matters and related issues, not forgetting the documentation for the whole process, which would facilitate enormously the task of producing a Methodological report after the census. This report should contain a description of all the census processes and tasks, timing, persons involved, coding procedures, editing rules, percentages of imputation, all the internal documentation produced, etc.

10.26 There is a need to increase ESO's power on Information Technology to face the challenges that the census will pose. Historically, in almost every country, census has a byproduct of setting the pace for future surveys on methods and technology. This will be true for Cayman only if the know-how could be transferred by training of ESO's personnel, which means that it has to have personnel to be trained. Good hardware and software are not enough if we do not have "peopleware", in quality and quantity, to handle them. One IT person (preferable two, since "if you have one you might have none very fast") is needed to start learning CSPro and PC-Axis (and also SPSS). A background on computer science is desirable, knowing a computer programming language like C++, C#, Pascal, Java or Delphi, is even better, and having some experience with macros for Microsoft Office (especially for Excel) would help. There is no need to have lots of experience, since it would be hard to find and very expensive to pay and maintain at ESO.

10.27 The scanners will need operators, as in 10.12, and people to handle the verification, as in 10.13, the exact number will depend basically on the overall schedule of the operation, everything must go in parallel, the scanners, the recognition, and the verification stations.

Capacity building

10.28 Software training is a must. All statistical officers at ESO should be able to work directly with their own data files, using a common "language" decided as the standard, for example, SPSS. It is a good choice, there is already some knowledge,

it is a matter of enforcing people to use it, or to create the need to do it. By having a "standard statistical language", and the statistical officers being proficient on it, ESO would be much more independent of external consultancies than they are now. Of course, for specific needs such as, for example, programming the data entry for the scanners using TELEFORM, it is difficult to think about knowing it profoundly, but one or two people should know it enough to be able to do some small modifications in the original design.

10.29 Basic training for the IT personnel must be concentrated first in CSPro. The US Census Bureau has summer workshops on CSPro, or a local training program could be arranged. SPSS has to be taught also, which I think would not be difficult to organize, and for PC-AXIS, a one week crash course should be enough to establish a neat website using aggregated tables. After discussing the matter with PC-AXIS sources (see Appendix III for more details), they agreed to program a little example using the Labor Force Survey tables, and also use it as the training material for the course.

Data dissemination

10.30 By looking at the ESO site at the internet, it handles only links to documents (most of them Acrobat PDFs), and links to other sites, mainly local, and some abroad. There are no data in tabular form (such as Excel or similar), so, it is difficult for the end user to do any extra calculations with the data in a PDF format, he/she would need to digitize the data again to be able to work with it. As a suggestion, besides the data in tabular format, it should also contain maps, much more graphics, and data base consultation procedures, through the use of PC-AXIS and REDATAM. PC-AXIS would take care of data dissemination of aggregated data, meanwhile REDATAM handles microdata. Both can be used very efficiently in the internet, by outside users.

10.31 One thing that has some appeal and could be easily implemented in the internet site is a Popclock (Population clock), containing the projected population estimates updated by the minute, although it has to be noted that the population projections data has to be available for the algorithm to be implemented. Several countries have a gadget like that, such as Brazil, New Zealand, the United States, etc. The US Census Bureau has one which is fully documented, and has also a clock for the World Population.

Documentation & Back up

10.32 Documentation is always under evaluated in a census project. It cannot be relegated to the final stages, were people want to finish and do a nice project wrap up presenting a beautiful document but meaningless and useless. It has to start at the early project stages, beginning with the project approval, time schedule, and a description of all its phases. Every step has to be fully documented, and these documents must be made available internally at ESO, and some of them externally also. The ideal situation would be having the Assistant Census Manager to keep all the documents, organized by topics, and at the end of the project, use them combined to help produce the Methodology Report.

10.33 Backing up is another issue that is always mentioned but rarely followed, and the examples of not having them are everywhere to be seen (see 1999 Population Census). During Census operations, a back up copy of all the files must be taken, before and after each step, mainly the imputation phase. More than one copy of the scanned data, before imputation, and of the imputed data, must be saved, if possible, elsewhere (in another government building) for natural disasters, floods, hurricanes, etc., prevention.

11. Results:

The results are commented in the previous item.

12. Contributions to the Country Program:

N/A

13. Value:

N/A

14. Recommendations

This is an abstract of the recommendations, accomplishments and results of this mission, already described in detail before. Of all of them, I consider the need for training the most important one for ESO, otherwise statistical processing at Cayman will depend forever on external consultancies. A concentrated effort on the development of their human resources should be emphasized. ESO needs an extensive assistance in IT-training.

14.1 It is very important to use the 1999 census microdata database, either in REDATAM format or otherwise. This data should be available to help in the analysis of the 2010 census.

14.2 ESO has the firm intention of using scanners for the 2010 census processing, but it would be too risky to do that with only one scanner in house, because if this one breaks or has any other problem during census processing, there is no back up. ESO needs to have at least two, preferably the same brand and model.

14.3 ESO needs to have as much independency as possible from outside resources, as far as IT is concerned. The way to do that is by using already tested software, and to train its personnel on its usage. Even though the mentioned software are very easy to use, there are times that an IT person is fundamental for the institution, to absorb the technology being transferred from several external sources of information (like consultants, training courses, etc). Even if ESO uses outside resources for some specific process, it sure will need some minor adjustments later, and ESO would not be able to wait until that resource would be available again (plus the costs involved). One IT person is necessary (preferably two) to start to concentrate all the ESO's knowledge on the matter.

14.4 There are a couple of software items that should be used for census processing, and later for other surveys as well. They are: CSPro, SPSS (or REDATAM), and PC-Axis, and in a much lesser extent, Microsoft Project. CSPro is an integrated software for data entry, editing and tabulation. SPSS is a complete statistical package, and REDATAM is for microdata database storage. It can be used from the very beginning stages of a census or survey, producing fast analysis and tabulations, up to the very end, for dissemination, through the internet. PC-Axis is for aggregated data (tables), which also produces maps and graphics, and can be used also through the internet. Microsoft Project is used for project scheduling.

14.5 ESO personnel need to be trained on all these software. Enough time has to be set aside for ESO staff to participate in this extensive IT-training.

14.6 The appointment of an Assistant Census Manager to be the census focal point person, and to work 100% with the population census.

14.7 Backing up all the census files before and after every census stage is a must.

14.8 Documentation is also a key issue in a census. To produce a Methodological Census Report at the end of the process should be one of the major goals of the Census Project.

15. Follow-up Action:

15a. Action	15b. Personnel Responsible	15c. Due Date
N/A	N/A	N/A

16. Lessons learnt and challenges

N/A

17. Contacts (Identification of experts/institutions (nationals or sub-regionals)

17a. Name	17b. Post	17c. Institution	17d. e-mail
Maria Zingapan	Director	ESO	
Ronnie Andersson	Chief Statistician		
Nicole Emmanuel-Jones	Census & Survey Supervisor		
Juliette Wright	Statistical Officer		
Wesley Howell	Computer Service	Government	

18. Appendices: (i) ; (ii) ; (iii) ; (iv) ; (v) ; etc.

Appendix I – Terms of Reference (TOR) for this Mission
Appendix II – Links

Appendix III – PC-AXIS Communications and Training program
Appendix IV – 2010 Census Main Planning Chart using Microsoft Project
Appendix V – Census Processing Main Flow
Appendix VI – Edit & Imputation Phase

19. Distribution

The findings were presented to and discussed with ESO personnel on Friday, May 23. The draft mission report was given to ESO in electronic format on the 24th of May.

Appendix I – Terms of Reference (TOR) for this Mission

TERMS OF REFERENCE

for the services of a

Consultant

To Provide Advice on IT-issues
in the Cayman Islands

PURPOSE

To conduct an assessment and provide advice on IT issues including recommendations on how to improve the IT system presently used by the Economic and Statistics Office (ESO), especially with reference to the upcoming Population and Housing Census. The census is scheduled for October 2010.

PROFILE OF CONSULTANT

Education
Post graduate degree in Statistics and/or IT.

Work experience
At least ten (10) years working in a government, regional or international statistical office, five years of which as a Senior IT specialist.

Other skills

Consulting experience in IT issues and expert knowledge of IT use in population and housing censuses.

WORK SETTING AND WORK PLAN

Under the direct supervision of Mr. Ronnie Andersson, Chief Statistician, the CONSULTANT will:

1) Undertake an assessment of the IT resources currently in place at ESO in the Cayman Islands and also assess the need for an in-house IT-expert within ESO;
2) Provide advice on the IT-system (hardware, software) for the 2010 Population and Housing Census, including dissemination;
3) Advice on the training needs in IT especially for the 2010 Population and Housing Census.

The Consultant will hold discussions with the Economics and Statistics Office and other relevant departments to assess the country's overall baseline capacity in IT. A report shall be submitted by the Consultant to ESO which shall describe the IT structure used so far and how it can be improved. The key resources, personnel, equipment, training, that need to be addressed in the improvement of the IT system should be described.

The assessment shall be followed by a meeting with the professional staff at ESO. The Consultant will prepare the materials for and conduct the meeting with the objective of a) disseminating the findings of the assessment, and b) presenting options and recommendations to address the findings of the assessment based on international or regional best practices.

The Consultant shall provide technical advice to the ESO in developing an IT-system to be used in the 2010 census. The proposed IT system must be tailor-made for the situation in Cayman Islands. The main objective is to

create an IT system which is suitable for the 2010 census but as the census is a capacity building exercise also for a more general capacity building in IT within ESO.

It should be noted that scanning (using TeleForm software) will be used in the 2010 census and ESO plan to use GIS both in mapping the EAs and also for dissemination.

DURATION

The assessment will require four (4) working days and the meeting, including preparations, should take one (1) day. The Consultant will provide advice, in a written report to ESO, on the IT-system for the 2010 census. In total the mission to Cayman Islands will be for 5 working days 20 to 25 May 2008. The report shall be submitted to ESO not later than June 15.

Appendix II – Links

Data recognition	www.rerecognition.com
Active software	www.active.com.uy
Top Image Systems	www.TopImageSystems.com
CSPro	www.census.gov/ipc/www/cspro
REDATAM	www.eclac.cl/redatam
PC-AXIS	www.pc-axis.scb.se
CANCEIS	www.stats.govt.nz/research-reports/evaluation-canadian-census.htm
MICROSOFT PROJECT	office.microsoft.com/en-us/project/default.aspx
BLAISE	www.cbs.nl/en-GB/menu/informatie/onderzoekers/blaise-software/default.htm?wbc_purpose=basicpublmethoden
Environmental System Research Institute (ESRI)	www.esri.com

Appendix III – PC-AXIS Communications and Training program

These notes are the results of a series of communications by e-mail with Lars Nordback (lars.nordback@scb.se), PC-AXIS manager at Statistics Sweden (SCB)

Costs

PC-AXIS price	US$ 4,400 per year, and a 50% more for the SQL module (which is not needed)
Training costs	32.000 SEK (US$5,000) for a week, plus all travel costs, accommodation, per diem for one person.

Software and Training details

1. Question: Is the SQL module needed?
   Answer: According to Lars, "SQL is not needed. There are possibilities to make PC-Axis files using PX-Edit (very easy to use from Statistics Finland) or from SAS (A module made by Statistics Finland), from Supercross (Australian product), TPL from the US company QQQ. I had a hope that the Redatam should make such a export function, but I have not seen anything yet. Some organizations have made PC-Axis file export functions by themselves from their internal database systems".

2. Question: How long should be the training?
   Answer: Training should take a week.

3. Question: How many people should be trained?
   Answer: Lars said "The number of people depends on the turnaround of staff, but I would suggest 4-5 persons".

4. Question: What should be their computer background?
   Answer: The background can be described like this related to the different part of skills:
   A: To make PC-Axis files using PX-Edit implies knowledge on the statistics to be able to form the best matrices. Education for these people is PC-Axis main module and PX-Edit.
   B: To set up the PX-Web website implies to have ideas on the customization to the present website and contact with the webmaster to link it into the website. Contact with IT concerning the routines for update of the website with new PC-Axis files and list of contents.

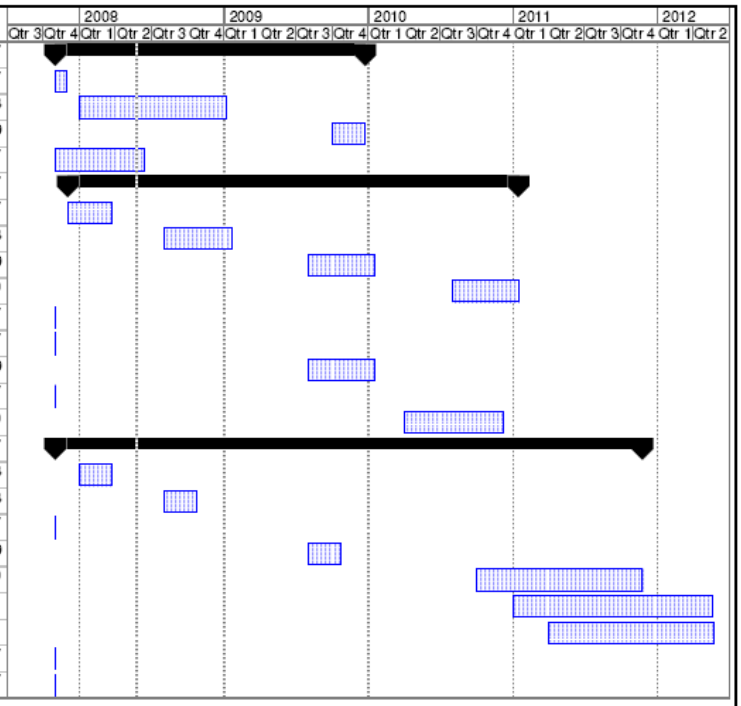
5. Question: Any previous knowledge of a programming language required?
   Answer: No programming language is required.

Training program

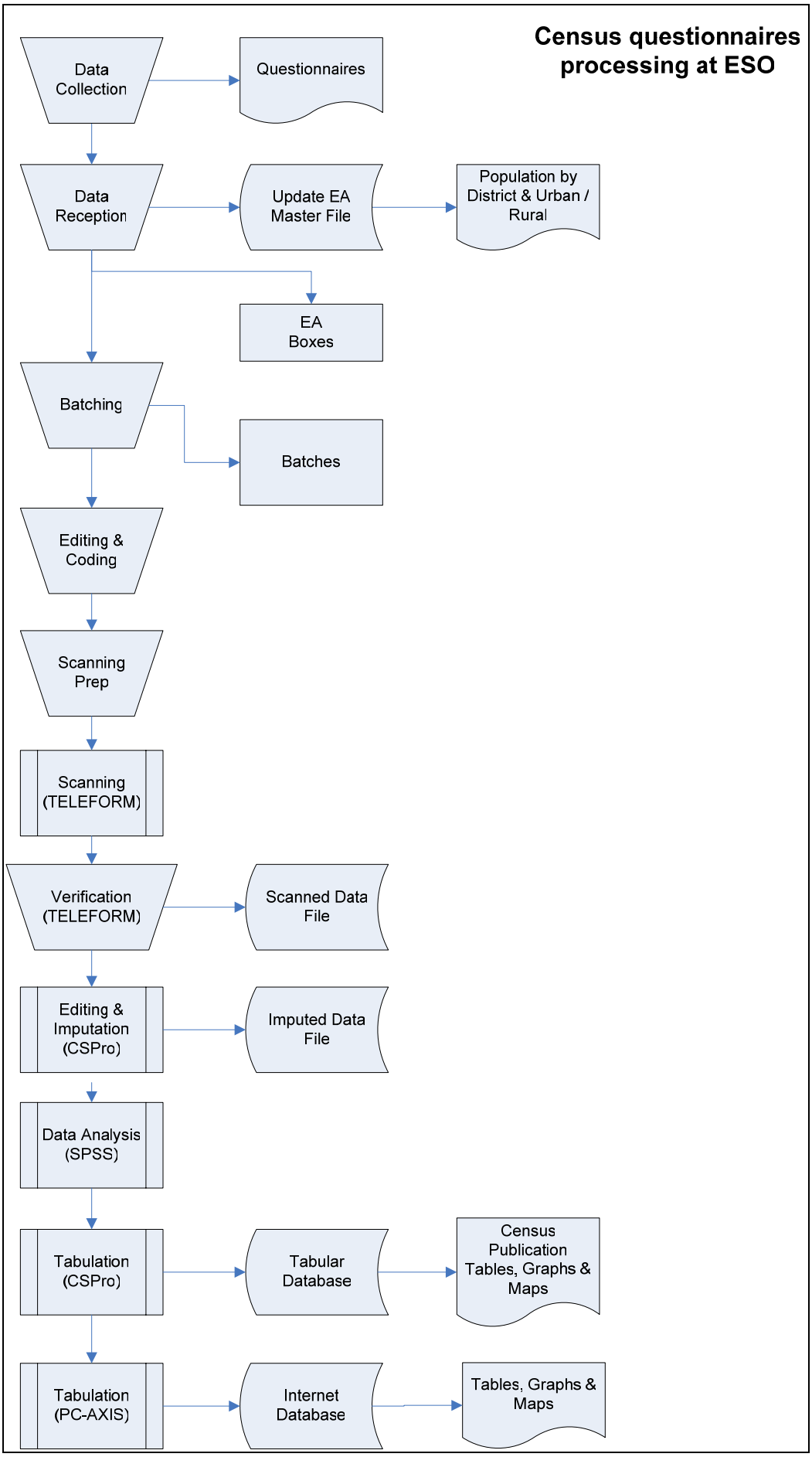
a. First day - A presentation of the concept including some demonstrations of PC-Axis main module, PX-Web and PX-Edit.
b. Late first day - Training course in PC-Axis main module.
c. Second day - Training course in PX-Edit and contacts with people with the Web and IT for initial preparation for the PX-Web (requires MS-IIS version 5 or higher)
d. Third day - Making PC-Axis files using PX-Edit in the morning
e. Third day afternoon - Start to customize PX-Web
f. Fourth day - Customizing PX-Web and making PC-Axis files
g. Fifth day - Presentation of the office new PX-Web site.

Appendix IV – 2010 Census Main Planning Chart using Microsoft Project

ID	Task Name	Duration	Start
1	**Legislation**	**554 days**	**Thu 1/11/07**
2	Legislation 1	20 days	Thu 1/11/07
3	Legislation 2	260 days	Tue 1/1/08
4	Legislation 3	60 days	Thu 1/10/09
5	Census Organization	160 days	Thu 1/11/07
6	**Budget / Procurement**	**809 days**	**Mon 3/12/07**
7	Budget 1	80 days	Mon 3/12/07
8	Budget 2	120 days	Fri 1/8/08
9	Budget 3	120 days	Mon 3/8/09
10	Budget 4	120 days	Mon 2/8/10
11	Mapping (HHREG)	1 day?	Thu 1/11/07
12	Questionnaire/Consultations	1 day?	Thu 1/11/07
13	Pilot Census	120 days	Mon 3/8/09
14	Publicity	1 day?	Thu 1/11/07
15	Training & Enumeration	180 days	Thu 1/4/10
16	**Data entry & Processing**	**1055 days?**	**Thu 1/11/07**
17	Consultant	60 days	Tue 1/1/08
18	Consultant	60 days	Fri 1/8/08
19	Programming	1 day?	Thu 1/11/07
20	Test	60 days	Mon 3/8/09
21	Data entry processing	300 days	Fri 1/10/10
22	Analysis, dissemination	360 days	Mon 3/1/11
23	Evaluation	300 days	Fri 1/4/11
24	Staff training	1 day?	Thu 1/11/07
25	Management/Planning	1 day?	Thu 1/11/07



Appendix V – Census Processing Main Flow



Appendix VI – Edit & Imputation Phase