USES AND PRACTICES OF COMPUTER APPLICATION IN DEMOGRAPHIC RESEARCH REGARDING NEPAL

Govinda Prasad Adhikari*

Abstract

The computer was used to model, simulate and analyze social phenomena in the research. This study aims to analyze the uses of computer software in demographic research in Nepal. Information was gathered from secondary sources from reports of GOs and NGOs, manuals and guidelines of various software packages. Informationwas taken from the United Nations software package, World Health Organization, United Nations Population Funds and software package of the U.S. Census Bureau. Other's information was taken from the National Housing and Census and Demographic Health Survey of Nepal. To achieve the objective of this study data are analysed in a descriptive method. This study has focused on computer software mainly SPSS, Mort Park, CS Pro, and STATA.

The roles of computer have been helpful in quantitative research. Its role is also important when the research is on a large sample. The computer assists the researcher throughout different phases in the quantitative research process. The major five phases of the research process are the conceptual phase, design and planning phase, empirical phase, analytic phase, and dissemination phase. The use of computer software is helpful in data entry, data editing, and data management. The data analysis software is supportive of processing and manipulating information and analyzing the relationship between variables. Therefore, the use of computer software can save money, time, and resources in research. In the case of Nepal, the Integrated Microcomputer Processing System (IMPS) was in the census in 1991. STATA and SPSS software packages were used for data editing, management, and tabulation of the census results. Open Data Kit platform and CAPI was used for data collection and the CS Pro software was used for data management in various demographic studies in Nepal.

Keywords: Computer software, census, survey, demography, research

Introduction

Demography is the statistical study of the population. Formal demography covers the measurement of the population process. Population Studies analyzes the relationship between social phenomena. Therefore, the population study covers all phenomena of social and cultural factors. It encompasses the examination of the social consequences of demographic change. The three main variables underlying population change are fertility, mortality, and migration. These are associated with factors such as age at marriage, the

*Mr. Adhikari is an assistant professor of Population Studies, Padma Kanya Multiple Campus, TU, Nepal.

proportions of marrying, contraceptive use, levels and types of morbidity, rural-urban migration, and so forth. Social demographers are those who seek to understand the processes of social factors such as the levels and distribution of income, levels of education, and the position of women, religion, and economic development. The possible linkages between variables are usually studied by social surveys (Duncan, 2008).

In 1948 the U. S. Bureau of the Census produced the first commercial electronic computer. In the early 1960s, the first book was devoted entirely to computer applications in social science. The book has focused on how to apply computer in designing and developing new software. During the 1980s, universities and colleges began to acquire microcomputers. The American Council of Learned Societies (ACLS) survey of 1985 reported that 50 percent of sociologists had a computer for exclusive use (Morton and Price 1986). A survey of academic departments supported by the American Sociological Association found that slightly less than half of the sociology faculty reported having immediate access to microcomputers (Computer Applications in Sociology; Koppel, Dowdall, and Shostak, 1985)

Computational social science refers to the academic sub-disciplines concerned with computational approaches to the social sciences. This means that computers are used to model, simulate, and analyze social phenomena. Fields include computational economics, computational sociology, culturomic, and the automated analysis of contents, in social and traditional media. It focuses on investigating social and behavioural relationships and interactions through social simulation, modelling, network analysis, and media analysis.

Two terminologies relate to each other: Social Science Computing (SSC) and Computational Social Science (CSS). In literature, CSS is referred to as the field of social science that uses computational approaches in studying social phenomena. On the other hand, SSC is the field in which computational methodologies are created to assist in explanations of social phenomena. Computational social science revolutionizes both fundamental legs of the scientific method: empirical research, especially through big data, by analyzing the digital footprint left behind through social online activities; and scientific theory, especially through computer simulation model building through social simulation. It is a multi-disciplinary and integrated approach to social survey focusing on information processing using advanced information technology. The computational tasks include the analysis of social networks, social geographic systems, social media content, and traditional media content (www. wiki/Computational_social_science).

In the field of demographic research, most of the researchers focus on a quantitative approach. Most of them are empirical in nature. They are focusing on big data, and it needs to integrate into social simulation. So, to simulate the big data into a social issue, the

researcher needs advanced computer softwares. This paper tries to analyse the uses of computer software applications to simulate big data into the social issues in the demographic survey of Nepal.

Objective of the Study

The objective of this study is to analyze the use of computer softwares in demographic research with special reference to Nepal.

Data and Methods

This study mainly used secondary sources of information. Information is gathered from reports of GOs and NGOs, manual and guidelines of various software packages. This article mainly focuses on the United Nations software package: World Health Organization (WHO), United Nations Population Funds (UNFPA) and public domain data processing software package of U.S. Census Bureau. Information is used from National Housing and Census Report of Nepal, Demographic Health Survey of Nepal and their use of software in data simulations process.

To achieve the objective, secondary informations are analysed in descriptive method. This study has focused on the computer softwares mainly SPSS, Mort Park, CS Pro, and STATA which are mainly used for data entry, data tabulation, and to perform various statistical methods.

Results and Discussion

There are a multitude of software programs designed for use with quantitative data analysis. In quantitative research, statistical analysis is common. Social scientists are using such software in social science researches. Most of the software based on statistical package applications are in the social sciences. The statistical software for quantitative analysis generates the table, chart, and plots of distributions and trends for the report. On the other hand, the computer software package for qualitative data analysis helps with text retrieval, coding, and building conceptual networks in qualitative research.

Computer is indispensable throughout the quantitative research process. The role of a computer becomes more important when the research is on a large sample. Data can be stored in computers for immediate use or can be stored in auxiliary memories like floppy discs, compact discs, universal serial buses (pen drives), or memory cards so that the same can be retrieved later. The computers assist the researcher throughout different phases in the quantitative research processes. There are five major phases of the research process: the conceptual phase, design, and planning phase, empirical phase, analytic phase, and dissemination phase.

The conceptual phase consists of the formulation of a research problem, review of literature, theoretical framework, and formulation of hypothesis. Computer helps for searching the literature and bibliographic references stored in the electronic databases of the world wide webs. It can be used for searching the literature in the form of books, journals, and other newsletters at the libraries which consume a considerable amount of time and effort.

The design and planning phase consists of research design, population, research variables, sampling plan, reviewing research plan, and pilot study. The role of the computer for sample size calculation is very important. Most common software (like NCSS-PASS, GESS) is available to calculate the sample size required for a proposed study. The researcher has also used various softwares (like Google Maps, Google Earth, and Geographical Positioning System) to find out the location and sample unit in the research.

The empirical phase consists of collecting and preparing the data for the data analysis. In research studies, the preparation and input of data is the most labour-intensive and time-consuming aspect of the work. Data gathering can be carried out in many ways. Among them, Computer Assisted Personal Interviewing (CAPI) and Paper and Pencil Interviewing (PAPI) methods are common and the most popular methods for primary data collection. PAPI is a more traditional method that involves questionnaires or surveys in paper form whereas CAPI is a more popular method that involves tools like tablets, smartphones, or laptops to record responses (https://scm.ncsu.edu). Typically, the data will be initially recorded on a questionnaire or record form suitable for its acceptance by the computer. To do this, the researcher, in conjunction with the statistician and the programmer, will convert the data into a Microsoft word file or excel spreadsheet. These spreadsheets can directly be opened with statistical software for analysis.

Data Storage is such a space where the data are obtained from the subjects and stored in computers as word files or excel spreadsheets. This has the advantage of making necessary corrections or editing the whole layout of the tables if needed, which is impossible or time-consuming in the case of writing papers. Thus, computers help in data entry, data editing, data management including follow-up actions and so on. Computers also allow for greater flexibility in recording the data while they are collected as well as greater ease during the analysis of these data.

The data analysis phase consists of statistical analysis of the data and interpretation of results. In data analysis, ample software (like SPSS, MortPak, CS Pro, and STATA) is available to perform the various statistical methods. These data analysis softwares are used to process and manipulate information, analyze the relationship and correlation between the dataset by providing quality analysis, content analysis, and decision-making methods. The

data analysis software is classified as exploratory data analysis and confirmatory data analysis (www.educba.com). Therefore, a computer is useful for statistical analysis to monitor the accuracy and completeness of the data.

The research dissemination phase is the publication of the research study. A computer is used in research publishing to type in word format and to convert into a portable data format (PDF) and store and/or publish on the World Wide Web.

Uses of Computer Softwares in Demographic Research

Open-source software is available with a license. The holder of the copyright for the software has the right to distribute the software. The researcher should be aware of the packages of open-source software. Some open-source software packages developed by the staff of the U.S. Census Bureau are available for free. These programs are used by researchers for data collection, data processing, and analysis. Some softwares mostly used in demographic research are discussed below:

Census and Survey Processing System (CS Pro)

The Census and Survey Processing System is a public domain data processing software package developed by the U.S. Census Bureau and ICF International. The main purpose of this software framework is to design applications for data collection and processing. It is a public domain software package used for entry, editing, tabulating, and disseminating census and survey data. It is designed as a user-friendly application. It supports data collection on Android devices (phones and tablets). It supports smart data transfer from android or windows devices to a server running CS Web. CS Pro is helpful for quantitative research to create data files, for analysing simple spatial maps and create a table for information dissemination (www.census.gov).

The United Nations Software Package for Mortality Measurement (MORTPAK)

MORTPAK is a software package for demographic measurement in developing countries with a special emphasis on mortality measurement. Version 4.3 is now available for free download using the link to the left. The present volume contains the working manual for MORTPAK for Windows, the United Nations software package for demographic measurement in developing countries. Version 4.3 of MORTPAK brings twenty varieties of applications in the areas of population projection, life table and stable population construction, graduation of mortality data, indirect mortality estimation, indirect fertility estimation, and other indirect procedures for evaluating age distributions and the completeness of censuses. These packages incorporate techniques that take advantage of the United Nations model life tables and generalized stable population equations. (www.mortpak_manual.pdf).

Statistical Package for the Social Sciences (SPSS)

Statistical Package for the Social Sciences (SPSS) is used for statistical analysis in quantitative research. It was acquired by IBM in 2009 and the current version (2015) is named IBM SPSS Statistics. The software is now popular in the field of health sciences and marketing. The IBM SPSS software platform provides advanced statistical analysis, text analysis, open-source extensibility, integration with big data, and seamless deployment into applications. With the help of SPSS software, the researcher can use statistical analysis and hypothesis testing approaches (www.ibm.com).

The Epi Data

Epi Data is a group of applications that are used for the creation of documented data structures and analysis of quantitative data in quantitative research. It is widely used to create and analyze large amounts of data. The World Health Organization (WHO) uses Epi Data in its STEPS method of collecting epidemiological, medical, and public health data, for biostatistics, and other quantitative-based projects (https://epidata.dk/index.htm).

STATA

STATA is a general-purpose statistical software package that was created in 1985 by Stata Corp. It is used in data management, statistical analysis, graphics, simulations, regression, and custom programming. Most of its users are in the fields of economics, sociology, political science, biomedicine, and epidemiology. It also has a system to disseminate user-written programs that lets it grow continuously (Stata Corp, 2014).

Word Processing Software

Word processing software is used to organize a text document, such as a resume or a report. Typically entering a text by typing and the software gives a tool for copying, deleting, and various types of formatting. Most of the writers have written their documents by ink before inventing the word processing software. Computer technology has invented word processing software. Word processing includes several tools which are helpful to format pages, organize text into columns, add page numbers, insert illustrations, etc. Text editors have a very different purpose from word processing software. They are used to work with files in plain text format. Functions of word processing software are a) creating, editing, saving, and printing documents, b) copying, pasting, moving, and deleting text within a document, c) formatting text such as font type, bolding, underlining, or initializing, d) creating and editing tables, e) inserting elements from other softwares such as illustrating or photographs and f) correcting spelling and grammar (What Is Word Processing Software? Definition, Types & Examples, 2013)

Demographic Analysis & Population Projection System (DAPPS) Software

DAPPS is a program designed to help users analyze and produce population projections with ease. It accomplishes this through a user-friendly spreadsheet interface for data entry and the projection power of RUP. For population projection, it requires at least three inputs: base population by age and sex, mortality structure by age and sex, and a fertility pattern by age of mother (www.census.gov).

Population Analysis System (PAS) Software

PAS is a set of Microsoft Excel workbooks developed by the U.S. Census Bureau containing frequently used procedures and methods in basic demographic analysis. It helps facilitate the analysis of available data for the following topics: age structure, mortality, fertility, migration, distribution of population, urbanization, and population projections. The workbooks are distributed with two manuals describing the demographic methods they implement and the procedures they perform. The PAS is a collection of Microsoft Excel workbooks for population analysis named Population Analysis with Microcomputers volume one and Population Analysis with Microcomputers volume two. These two volumes provide the analytical techniques and instructions of PAS software (www.census.gov)

Rural-Urban Projection (RUP) Software

RUP is a fully developed demographic projection program for preparing cohort-component projections for a population defined either in five- or single-year age groups one year at a time. The cohort component method projects each age and sex cohort over time based on the components of growth. Annual births create new cohorts while existing cohorts are decreased by mortality and either increased or decreased by migration. The RUP program has features that allow a considerable amount of flexibility for specifying fertility, mortality, and migration. As a result, RUP can be used to produce estimates for years where data on these components are available followed by projections into the future. RUP also includes a wealth of output options that allow a detailed examination of the results.

RUPAGG can calculate the third area as the sum of the two areas. This flexibility allows a demographer to create a projection model that accurately reflects what is known about the demographic situation in a country, making maximum use of available data. The RUP combined programs allow modelling of population changes due to catastrophic events (e.g., famine or civil conflict) when they are confined to one-half of a calendar year. The RUPEX Excel interface allows the user to select a RUP, RUPAGG, or RUP Combine input file from a user-defined subdirectory, edit the file, create a duplicate file, print the file, run the projection, and specify and generate output tables and charts for the projection's base year and any projection years. The graphics capability of the RUPEX interface does not provide

diagnostic tools previously available to RUP users; it allows the user to more easily check the plausibility of assumptions, and to compare projection inputs and outputs dynamically over time (www.census.gov.html).

Maptitude Demographic Mapping Software

Since 1995 Maptitude mapping software has made it easy to access a wealth of Census data for entire countries. It can provide a map library that contains numerous pre-designed maps listed by Census category. These include heat maps of demographic, income, and housing data. In addition to Census sources, Maptitude provides extensive non-Census information. It provides people in a particular age and income group. Maptitude Geographic Information System (GIS) software gives you the tools, maps, and demographic data you need to analyze and understand how geography affects you and your business. Maptitude has the richest feature set and highest performance of any mapping software. Maptitude offers comprehensive desktop mapping software functionality and provides many new and enhanced features that make it easier for you to create and edit maps (www.caliper.com).

Household Registration System (HRS)

The Household Registration System (HRS) is a software system currently in use in eleven research sites in Africa and Asia. The HRS exploits recent developments in object-oriented programming and automated program generation to simplify the process of developing data management systems for a diverse collection of longitudinal household studies. The computational foundation of the HRS is a relational database system that resolves many of the complex data management issues associated with monitoring births, deaths, marriages, and migrations in a fixed geographical area. It is working in the research of testing the efficacy of malaria vaccines, assessing HIV epidemiology, researching pneumococcal pneumonia epidemiology and prevention, examining cerebral spinal meningitis squealed, testing family planning impact, and exploring options for health care reform. The HRS FoxPro code is available on the Population Council website (Phillips, MacLeod & Pence, 2000).

Uses of Computer Software at National Census and Survey in Nepal

In the National Demographic and Health Survey (NDHS) 2006, the listing was performed by a team of listing and mapping persons manually by using The Global Positioning System. GPS is a satellite-based radio-navigation system that provides geolocation and time information anywhere on or near the earth from the satellites. Data processing was performed by using an MS computer (MOHP, New ERA, and ICF International Inc., 2007). The NDHS 2011 used CS Pro software for data entry at ASUS Eee T101MT tablet PCs. Code division multiple access (CDMA) wireless technology via Internet File Streaming

System (IFSS) was used to transfer data from the field to the central office. This package was tested for the first time in Nepal (MOHP, New ERA, and ICF International Inc. 2012).

In the NDHS 2016, data were collected by using the CS Pro software through a tablet PC. All electronic data files were transferred via the IFSS to the New ERA central office in Kathmandu. These data files were registered and checked for inconsistencies, incompleteness, and outliers. The biomarker paper questionnaires were compared with the electronic data files to check for any inconsistencies in data entry. Data entry and editing were carried out using the CS Pro software package (MOH, New ERA, and ICF International Inc., 2017).

There is not a long history of computer in Nepal. Nepal hired some Facit calculators for census in 1961. IBM 1401 a second-generation mainframe computer was used in 1971 census. In the 1981 census, CBS used a second-generation mainframe computer (ICL 2950/10) for data tabulation. The Integrated Microcomputer Processing System (IMPS) was used for data entry in the 1991 census. STATA and SPSS software packages were used for data editing, management, and tabulation of the census results in census 1991 and 2001. Both censuses used Microcomputers (like Pentium III for the data entry, Pentium IV & III for the data processing and tabulations, and high-speed laser printers) for data processing and report writing (CBS, 1991; CBS, 2001)

In census 2011, CS Pro software was used for data entry, editing, verification, and tabulation. CS Pro, SPSS, and STATA were used for data tabulation. Thereafter CBS has experience in the development and application of technological advancements in the current year. CBS firstly used a Tablet Computer to enumeration earthquake-affected housing surveys by using ODK (Open Data Kit) platform. The CAPI was used for the Nepal Labour force Survey 2018 and the multi-modal approach for Economic Census 2018. CBS has prepared a database for rural/municipal and EA maps for the forthcoming Population and Housing Census 2021 (CBS, 2018).

The CS Pro (Census and Survey Processing) package was used for data management in the Nepal Labour Force Survey (NLFS III) 2019. The use of tablets for data collection decreased the burden of data entry in this national-level survey. For the household survey section statistical package (STATA) was used. For a listing of the household and enumeration, an android-based program has been prepared in CS Pro. The information collected from the field using the main questionnaires had been received directly in the central server on time. It helps to minimize the burden of data entry, coding, and to reduce errors from various steps (CBS, 2019).

Discussion

The Microsoft Excel workbook was developed by the U.S. Census Bureau for basic demographic analysis. It was helpful to analysis of data for fertility, mortality, and migration, distribution of population, urbanization, and population projections. MORTPAK software was developed to measure mortality in developing countries. The MORTPAK applications work in the areas of population projection, life table and stable population construction, graduation of mortality data, indirect mortality estimation, indirect fertility estimation, and other indirect procedures for evaluating age distributions. CS Pro software is helpful to create data files, tabulation of information, analysis of simple spatial maps and dissemination of data. SPSS software is used for statistical analysis, integration with big data, and seamless deployment into applications. The EpiData is used for the creation of documented data structures and analysis of quantitative data in quantitative research. It is widely used to create and analyze large amounts of data. STATA is used in data management, statistical analysis, graphics, simulations, regression, and custom programming.

In the case of the national-level demographic census/survey of Nepal, the Integrated Microcomputer Processing System (IMPS) had been used from the census of 1991 to the data entry. STATA and SPSS software packages have been used for data editing, management, and tabulation of the census results. The Open Data Kit platform and CAPI has been used in the Nepal Labour force Survey 2018.

The Rural-Urban Projection (RUP) Software has been developed for the demographic projection (cohort-component projections) for a population in a defined area either in five- or single-year age groups. RUPAGG software has been used to calculate the third area as the sum of the two areas. The Maptitude mapping software is helpful to access a wealth of census data for entire countries. Maptitude GIS software gives the tools, maps, and demographic data. But the RUP, RUPAGG and Maptitude software are not currently used in the demographic study in the case of Nepal.

Conclusions

The use of computer software is very helpful in demographic research. It is useful for the calculation of sample size, to find out the location, and to find out sample units in the research. The computer software is helpful in the literature review to search various books, journals, and other newsletters. It is also useful for the data collection, data process, and analysis of research. The most useful computer software which was used in demographic research is: STATA, MORTPAK, CS Pro, Epi Data and R etc. Microsoft Excel was mostly useful for the basic demographic analysis. MORTPAK software was used to measure

mortality in developing countries. The MORTPAK application works in the areas of population projection and life table construction. The CS Pro software was used to create data files, summarization and analysis of data. The SPSS software was used for statistical analysis, integration with big data, and seamless deployment into applications. The EpiData was used for the creation of documented data structures and analysis of quantitative data in quantitative research. STATA is used in data management, statistical analysis, graphics, simulations, regression, and custom programming.

The Nepal Demographic Health Survey has been using SPSS and STATA software for data collection, data entry, data analysis since 2006 and CS Pro software was used for data entry since 2011. The census program of Nepal had used the IMPS software, the STATA and SPSS software for data processing in the census 1991 and 2001. The CBS was firstly used for the ODK program in the earthquake-affected housing survey. The CAPI program in the Nepal Labour force Survey 2018.Finally, various computer software like Rural-Urban Projection (RUP) for the cohort-component projections of population, RUPAGG software for calculation of the third area as the sum of the two areas and Maptitude software for GIS maps and demographic data are not currently used in Nepal for the demographic study.

References

Central Bureau of Statistics (CBS), Nepal. (1991). *National Population Census 1991.*

Government of Nepal, National Planning Commission Secretariat, Central Bureau of Statistics, Kathmandu.

Central Bureau of Statistics (CBS), Nepal. (2002)*. National Population Census 2001*.

Government ofNepal National Planning Commission Secretariat, Central Bureau of Statistics, Kathmandu.

Central Bureau of Statistics (CBS). (2019). *Report on the Nepal Labour Force Survey (NLFS III) 2017/18* Government of Nepal, National Planning Commission Secretariat, Central Bureau of Statistics, Kathmandu.

Central Bureau of Statistics (CBS). (2018). *Regional workshop on the use of electronic Data collection technologies in population and housing censuses, 24-26 January 2018, Bangkok, Thailand*. Government of Nepal, National Planning Commission Secretariat, Central Bureau of Statistics, Kathmandu.

Computer applications in sociology. Encyclopaedia of sociology

https://www.encyclopedia.com/social-sciences/encyclopedias-almanacs-transcripts-and-maps/computer-applications-sociology

Duncan G. J. (2008). *When to promote, and when to avoid, a population perspective. Demography*, *45* (4), 763–784. https://doi.org/10.1353/dem.0.0031

Koppel, R, G. Dowdall, and A. Shostak. (1985). *Using computers to teach sociology: A departmental survey. (na.).*

Ministry of Health and Population (MOHP) [Nepal], New ERA, and ICF International Inc.

(2007). *Nepal demographic and health survey 2006*. Kathmandu, Nepal: Ministry of Health and Population, New ERA, and ICF International, Calverton, Maryland.

Ministry of Health and Population (MOHP) [Nepal], New ERA, and ICF International Inc. (2012). *Nepal demographic and health survey 2011*. Kathmandu, Nepal: Ministry of Health and Population, New ERA, and ICF International, Calverton, Maryland.

Ministry of Health (MOH) [Nepal], New ERA, and ICF International Inc. (2017). *Nepal demographic and health survey 2016*. Kathmandu, Nepal: Ministry of Health

Phillips, J., MacLeod, B., & Pence, B. (2000). *The household registration system: Computer software for the rapid dissemination of demographic surveillance systems.* demographic research. http://www.jstor.org/stable/26348004

Mare, R. D., Maralani, V. (2006). *The Intergenerational effects of changes in women's educational attainments*. American Sociological Review. 2006

Stata Corp (2014). *NetCourse 101: Introduction to Stata*.StataCorp. Information about the course available from: http://www.stata.com/netcourse/intro-nc101/

What is word processing software? Definition, types & examples, November, 2013. https://study.com/academy/lesson/what-is-word-processing-software-definition-types-examples.html.

https://scm.ncsu.edu/scm-articles/article/data-collection-electronic-or-manual

https://www.educba.com/data-analysis-software

www.census.gov/data/software/pas.html

https://www.census.gov/data/software/cspro.html

https://www.census.gov/data/software/dapps.Download.html

https://www.census.gov/data/software/pas.html

https://www.caliper.com/maptitude/solutions/demographic-mapping-software.htm

 https://www.ibm.com/analytics/spss-statistics-software.

https://www.un.org/en/development/desa/population/publications/pdf/mortality/mortpak_manual.pdf