AN EMPIRICAL APPLICATION OF DATA MANAGEMENT TECHNIQUES IN ENVIRONMENTAL, EARTH AND SOCIAL SCIENCE RESEARCHES

By

Ifabiyi, I.P.
Dept. of Geography & Environmental Management, Fac. of Social Science, University of Ilorin, Ilorin

Author's Email: tokunifabiyi@yahoo.com ; tokunifabiyi@unilorin.edu.ng

ABSTRACT

Data is required in all facets of research endeavour and for it to be used and become useful; it must be scientifically collected, analyzed and interpreted. All these efforts involved highly technical procedures. This paper (i) discusses some methods of data management and (ii) exemplifies efforts involved in data management using primary sets of data. The example in this paper demonstrates procedures for handling large volume of data sets typical of social, environmental and earth science researches. In the empirical data applied to explain data management procedures, 58 variables were generated from a structured questionnaire used in soliciting information on rural water supply and climate change adaptation in Goronyo, Sokoto, state. The large size of variables of variables for the explanations of the relationships was reduced via the principal components analysis (PCA). The result of PCA was further processed with the multiple regression and multiple stepwise regression analyses; two methods that established relationships between quantity of water demanded and variables of climate change adaptation. The PCA and multiple regression analysis identified some explanatory variables that account for about 88% of variation in climate change adaptation and isolated indices of waiting time, water sourcing, water conservation, household size and occupation as useful in water resources management under climate change. The paper concluded that data is necessary in all facets of human lives; therefore, it must be collected using scientific procedures and must be properly interpreted using appropriate techniques. It is hoped that this paper will be relevant to a broad spectrum of researches especially in the social, environmental and earth sciences.

Keywords: Data management, environmental and social sciences, climate change, water adaptation.

INTRODUCTION

Data simply refers to information or attributes or sets of measurement about an object. Efforts at data collection dated back to antiquity. The first recorded attempt of data collection dated back to the bible Old Testament which contains instances of census in ancient Greek and Roman Empire. The first modern day evidence of data collection the effort of Sir John Sinchiar in his famous work on statistical Account of Scotland (1791-1799) and that of William The Conqueror in 1806 when he ordered the writing of the Doomsday,s Book. Since then, collection and collation of data has become an everyday use. It is required at household level by housewives normal upkeep of the home. For example, numbers of members of the household is required before budgeting for the family. In the barbers shop, the numbers of customers determine projections on his income and expenditure and it will be a guide on what extent of loan to seek for. More importantly a nation with dearth of data will no doubt be backward and

will find it difficult to plan and consequently fail in discharging her responsibilities to her citizens.

The aim of data collection is mainly to make explanation about different phenomena. The specific objectives of data collection include: aggregation and summary, measurement of distribution, analyses of different structures, and explanation of variation, testing of hypothesis, extrapolation and prediction. This paper will attempt a discussion of data collection, collation and manipulation efforts. Hence, the rest part of this paper is consequently devoted into doing this, by a consideration of data, qualities of a good data, sources of data, choices of data collection strategy and sample survey. Others are data analyses, error in data collection and problems facing data collection in Nigeria.

Types of Data

There are different types of data. Data can be classified in different ways. Croxton, et. al (1989) classified data into four types. These are discussed below.

i. Qualitative Data: These are largely data, which are descriptive. Such as information about sex of respondents, for example, whether single, married or divorce. These types of information are required for purposes of description and qualitative analyses of various objects of interest.

ii. Quantitative Data: The quantitative data are measurable attributes of different objects. They are most often numerical in nature example are level of income, height of an object, acreage of land devoted to irrigation, numbers of items produced, etc.

iii. Chronological Data: These are times series data, which show figures concerning a particular phenomenon at various specified time. Example are demographic attributes such as birth rate, age of respondents, monthly rainfall amount, number of goods sold per month, etc. chorological data are related to time.

iv. Geographical Data: These are geographical distribution, for example, the densities of population per state or number of NDE beneficiaries per Local Government Area, etc.

Data collected for any purpose must possess certain characteristics and where it did not meet up with this standard it must be treated with caution and may be discarded out rightly. According to Selltiz, et .al (1976), these qualities are four; It must serve a formulated research purpose. Its collection must be planned systematically. It must be recorded systematically and related to more general proposition than reflecting a set of interesting curiosities. It must be subjected to check and control on validity and reliability.Data according to Abumere (2001) are broadly from three major sources vis:

a. Recording of Transaction: This is the method of recording events as they occur. Examples of this include; information recorded on daily basis in vehicle registration centers including data on numbers and types of vehicles, information in airports and seaports such as incoming and outgoing of goods and people, information in private sector offices and departments such as records of daily transaction on production and volume of sales and many more. This type of data is largely inaccessible because bureaucratic bottlenecks.

b. Census : A census is a complete count of population, which could be made up of human beings, shops, farms, houses or industries. The count should include every member of the population of interest and there must not be any omission or double counting.
c. Sample Survey: Rather than enumerate individual as in the case of census, we might in fact enumerate a few or a fraction of the population using certain procedures to select a few. This is based on the principle of the law of statistical regularity than a set of objects taken at random from a larger group tends to produce the characteristics of large sample.
d. Apart from the aforementioned broad classification data can further be sourced from any or a combination of the following methods.
    i. Controlled Experiments: This involves the process of taking observations under different controlled conditions for the purpose of assessing effects of changes in the condition. This process involve; careful statement of the problem, choice of the features of interest in the population and choice of features affecting the variable of interest
    ii. Census: this entails total counting or enumeration of objects making a population in which is interested. They include population census, agricultural census, housing census, industrial population census and so on.
    iii. Sample Survey: This is an examination of a part of a population to make inferences about the population. This entails brief contact with a number of persons who are representatives of the population to be described. It provides reliable results at low cost. The result are adapted to statistical tools, all sample survey must have an objective. Sample survey may be divided into three : social survey: mostly done by government, market research: conducted by private firms and entrepreneur that are interested in projecting future of their sales, and public opinion poll: It is used for assessing political opinions, and public reactions to legislative acts among other uses
    iv. Administrative Activities: These are inform of social statistical registrations such as registration of marriages, births, deaths. These are mostly done by completion of forms, which are returned to Central Data Processing Units, a good example of these are government ministries and parastatal such as the Federal Office of Statistics and Ministry of Social Welfare where marriages are registered from time to time. (Adamu and Johnson, 1975).
    v. Participant Observation: This is common in anthropological studies when researchers are living-in, with the community under investigation. For example, to study crime in a place, the investigator may have to be resident in the area. This is based on the principle that after a reasonable length of time after which some level of integration might have taken place, people will be free to act naturally in the presence of the researcher and the necessary recording and data collection can be conducted (Olorunfemi 2004).
    vi. Archival and Non-Communicative Source: These are obtained without communication with respondents. It is common in demographic, archeological and socio-metric data collection process. This type of information has lower cost collection.

vii. Direct Observation: These are gathered by direct communication. Examples are interviewing or questioning of respondents. Others are observation of styles of houses, reactions of people to special events. Here the investigator seeks a means of detecting differences without being a disturbing influence.

viii. Archeological and Photo Interpretation: Data on past civilization could be obtained from the artifacts of yesteryears left behind by ancestors. Examples of these include ancient weapons, fossilized seeds, living quarters. Such artifacts can be dated by carbon-14 test. Anthropologist and archeologist can also reconstruct and interpret ancient photographs, etc. this approach is also useful in land use planning, mineral exploitation, etc

ix. Participatory Rural Appraisal **(PRA):** This is a new strategy of data collection which came out of a realization and awareness of the failure of conventional development approaches and data collection strategies in meeting the needs and resource management among poor people. This has led to the exploration of alternative methodologies of investigation. (Groot and Maarleveld, 2000; Chambers 1994; Slim and Mitchell, 1992). Some of the more developed and tested methods of PRA include the following:

a. Participatory mapping of settlement by the inhabitants including the plotting of important landmark such as rivers, roads, etc
b. Group Discussion: organized brainstorming workshop
c. Transect Walking: the researcher or planner walks with the people through an area of interest, observing, listening and asking questions.
d. Direct Observation: Do it yourself approach, people teaching the researcher the way they do it.
e. Force Field Analysis: A technique to visually identify and analyze forces affecting a problem and giving alternative solution through group work and activities.
f. Collective Identification of resources including access, management and control and including sources of income, health and links to other settlements.
g. Establishment of formal and informal groupings that can provide a focus for and maintain the momentum of community driven development.
h. Matrix scoring and ranking: A wide range of matrices can be drawn using local matrices and with opportunities for participant or group to change their opinions on issues.
i. Community costing: A technique by the community to choose community-felt need projects with participants estimating costs of the projects proposed within the community.

CHOICE OF METHOD OF DATA COLLECTION STRATEGY

Utmost care should be used in selecting what method to use in data collection as this can mar or make a research. However, Casley and Lury (1987) identified certain guidelines for choosing method of data collection. They are:

i. Assessment of Priorities: It must be established that the method chosen should be the best at his disposal and that equivalent information cannot be more cheaply.
ii. User-Research Dialogue: The two parties must communicate; this is to enable a universally accepted maxim.
iii. Definition of problem and the researcher objectives: The research objectives and research problem must be clearly defined.
iv. Existing Knowledge: The researcher should be alive to possible improvements, which may be obtained by exploring secondary or alternative sources of data.
v. Secondary Users: The researcher should be aware of the other people and institutions who will likely be interested in the results of the proposed study.
vi. When a researcher in problem- policy-oriented research, he should prepare for alternative remedies to the specific data collected.
vii. Resources and Time: The resources available for the research and time, the research will take should be discussed.
viii. Time and Cost Analysis: There should be a timetable and cost analysis of the research effort before the start of the exercise.
ix. Rapid Assessment: This is achieved through a proposal of data collection methods and timetable that will produce the information within policy dictated deadlines.

Sampling Procedure

Sampling survey remains the most versatile of the methods of data acquisition, this is because most often it become very difficult and almost impossible for an investigator to study every object in its target population and therefore would have to result to sampling. Hence, because of the important of sample survey in data collection efforts, the following section will be devoted to a further look at it. The principle underlying sampling is that a set of object taken at random group tends to reproduced the characteristics of the larger group. It has several advantages: It is cheaper, it consumes less time, more question are asked and more information are obtained, if properly done especially by experts better result are obtained.

Selecting the Sample

In order to avoid bias and ensure that the sample selected reflects the properties of the population; sample has to be selected following certain methods. Some of which are as follows:

i. Simple Random Sampling: This is a technique that aims at eliminating bias in the choice of respondents. This is by selecting ‘n’ (sample) units out of ‘N’ (population) such that every member of set of population or community has an equal chance of being selected or drawn. This also involves the use of random numbers. This is done without replacement . The advantage is that every item or person in the sampling frame has equal probability of inclusion. It is easy to use.

ii. Systematic Sampling:
It is convenient and simple. It is an alternative to the above, it has more even coverage. Items are picked at some regular intervals, for example every 10^{th} items on a list, every 20^{th} grid square, every 100^{th} line across a map, etc. It does not give equal opportunities to every member of the population to be selected

iii. Stratified Random Sampling: This is more scientific and rigorous. In this case, the population N is first divided into subpopulation of N1, N2, Nn units, based on some observed set of criteria, this is done without overlapping. Each of the subpopulation is called a stratum and samples are drawn from each stratum randomly or systematically. The variables used in stratification can be income, that is, low, medium or high.

However, various combinations of these three can also be used such as Stratified- random, stratified systematic sampling, systematic random sampling, etc.

iv. Others : Apart from the sampling methods discussed above several others are also available especially in specialized fields. Some of these are purposive sampling, random transect sampling in land use and land capability studies, quota sampling, random routes sampling, snowball sampling, transects sampling, quadrat sampling in vegetation and ecological surveys , random walks sampling which is used when an investigator is operating under a very limited time limit.

In general, it should be stressed that more than one or a combination of these methods can be used at a time; they can be combined, as this will offer better results. However the choice of these methods depends on the intension of the investigator and some of the factors earlier raised.

Data Analysis

The objectives of data analyses are many: to test hypothesis, to make inferences, for data description, prediction and projection/forecasting and classification purposes. Three major methods of data analysis are commonly adopted. These are descriptive, predictive and inferential methods.

i. Descriptive methods include all methods that can summarize or describe important features in a set of data without going any further to draw conclusions or inferences about a large group. Examples are frequency distribution, measure of central tendencies such as mean, median, mode. Others are dispersion methods such as standard deviation, variance, range and inter quartile range, co efficient of variation, etc. all the graphical methods such as all forms of charts and graphs also belong to this category

ii. Predictive statistics are used in solving problems that might arise in the future or rather knowing the best situations that might minimize costs while maximizing benefits.
This method is useful in many respects. This is the method that is normally used for problems of missing data. often a times investigators encountered problems of missing data and most time he cooks up these missing ones. But alas several methods are available for overcoming this problem. The methods commonly used are the moving averages, trend surface model, autogressive models, etc. (Thrift and Oaks, 1974;Tobler 1973;Tobler 1970)

iii. Inferential statistic**:** Inferential statistics are those that enable us to make generalization that go beyond the data. This is common in social sciences. Inferential statistic can be categorized into two, these are parametric and non-parametric methods. The parametric

methods data require the population from where the data is drawn to have certain conditions or characteristics. It must be normally distributed, linear, random, etc. Non parametric data on the other hand do not require these conditions. However the non-parametric tests are not as powerful as the parametric tests (Abumere, 2001).

All these have been facilitated with the coming of computer. The advent of computer has brought in lots convenience in data analysis and overall data management. A lot of computer programs and models are now available to do all these in twinkle of an eye. Examples of such programs are SPSS, SAS,EPI info, Excel and a lot more in different specialized fields.

Problems of Error in Data Management

According to Abumere (2001) two types of error are expected. These are random error and systematic error. Random error may appear mild. For example, asking people about their age, particularly women may under estimate their while villagers will over estimate since they thought some level of prestige accompany over estimation of age such error cancels out. However, in the other dimension questions on the level of income of respondents may prompt them to under estimate their income particularly if such work will be used to estimate tax.

Harvey (1969) has identified five types of error;

i. Observer Error: this emanate from inability to measure properly. It can even be due to incorrect calibration.
ii. Environmental Error: when condition changes with environmental changes, for example certain instruments react to temperature through and expansion.
iii. Error due to the observed; which is due to inherent variability of the thing to be studied, for example, people changes their views at any time.
iv. Error due to behavior of the observer: such as dresses, sex of interviewers, etc. respondents gives answer the way you want.
v. Condition effect; for example, if one house hold is interviewed today, and the adjacent house the following day, the interviewing conversation might conditioned the responses on the second interview.

Problem of Data Collection in Nigeria.

i. Political Will: several administration of government have not instituted any policy to make sure data collected at a scale relevant for research and policy making. Scarcity of data explains the reason for the collapse of government infrastructures; a condition which has brought untold hardship to the Nigeria populace.
ii. Capital and Resources**;** collection of data is an expensive business that needs a lot of capital in terms of liquid and capital infrastructures. However, the reward for data collection is so much but may not be directly physical. Hence, entrepreneur, government and people are reluctant in investing in this business. Thank God for World Bank and UNDP for co- finance Census 2006 in Nigeria.
iii. Problems of ignorance and illiteracy: most officers charged with the task of data generation do not know the importance of their job. Also, an average person even entrepreneur may not know the importance of data generation and collection. There is a bad attitude to data management in Nigeria.
iv. Bureaucracy: Researchers, including this presenter has been frustrated severally in the bid of collecting secondary data. Most of the time you need a lot of clearance,

identification and all sorts, which are unnecessary and frustrating. Some of this attitude is mostly deliberate by civil servants. The end results are poor planning as some data are simply cooked up.

EMPIRICAL STUDY: ***Rural Water Supply Adaptation in Climate Change Scenario in Goronyo Local Government Area of Sokoto.***

The example of human response to climate change in Goronyo is hereby used as a case in point. The data required in this study are mainly information on water use characteristics of the rural people, people's perception on climate change, and on the adaptation strategies of the local people. This information were solicited with the use of structured questionnaire. The questionnaire used in this study was divided into 4sections. Section A focused on primary information, Section B on water use characteristics, section C climate change perception and section D focused on response and adaption strategies in the local government area.

Questionnaire administration was done through systematic random sampling. Hence, all the 11 wards of the study area were covered. 15 questionnaires were administered in each ward. These make a total of 165 questionnaires in the whole of the study area. A total of 58 variables were generated from the administered questionnaire. These variables can be classified into 4 categories. These are: primary attributes, water use components, climate change perception and climate change response strategies.

In view of the size of these variables a Double Stage Reduced Model (DSRM) involving both factor analysis (FA) and principal component analyses (PCA) were used to remove redundancy from the variables. Factor analytical technique was first used to identify significant variables out of the 58 generated in this study. The result of this first stage of reduction showed that 17 factors were generated with a total of 30 highly significantly loaded variables (Table 1). The percentage of explanation in this case was 88%. Principal component analysis on the other hand, was used in the second stage reduction which involved the selection of the identified 30 variables into the principal component model. The result of the second stage showed that 10 components were responsible for the explanation with a total cumulative explanation of 80% to the variance (Table 2). These 10 components are the underlying factors of climate change and water resources in Goronyo L.G.A. of Sokoto state Nigeria.

Components of water use, climate change and adaption.

i. Component I: This factor is tagged household time**; it** is also strongly loaded on maintenance of water source. This factor contributes 15.2% to the explanation of the variance. The amount of time spent per head in collecting water is a signal to individual ease of getting water. In the study area, water sources are almost 100% from groundwater. Hence, it is available at the point of need. It maintain by individual. Hence, shorter time is spent collecting water. This component is an *index of waiting time*.
ii. Component II is tagged access to water. It has high loadings on 3 variables. These are access to water, reduction in agricultural output, and impact on cottage industries. It contributed 12.1% explanation. This component is an *index of water accessibility***.** Accessibility to water will affect water use in the agricultural sector and the cottage industries. If water is accessible then it will impact on the use of water in the farms and in the local industries.

iii. Component III is known as reduction in water use it is an adaptive strategy due to water scarcity induced climate change. This factor has significant loadings on beast of burden, water conservation, and reduction in water use. It is an *index of water use conservation.* It has contribution of 9.70% to the total explanation. Reduction in water use and water conservation coupled with the usage of donkeys in fetching water are ways by which the Goronyo people adapt to water shortage due to climate change. Donkey is a beast of burden that has capacities to carry about 120litres of water at a time. Hence the use of donkeys to collect water reduces the numbers of trips to water points and it is also more convenient compared to head portage system of fetching water. This has proofed a wonderful adaption mechanism for fetching water in the study area.

iv. Component I has high loadings on water borne diseases, sources of water, and distances to available water points. This factor has contribution of 9.13%. It is an *index of water sourcing.* The source of water is a signal to the level of water contamination. However, water contamination, particularly as referred to spread of water borne disease is not common phenomenon as the major source of water in Goronyo local government area is mainly hand dug wells and boreholes, more importantly, distances to water are generally short. This is expected in view of the fact that underground water is available at close to the source of supply, consequently, individual households have hand dug wells in their homes and compound.

v. Component V is contributing 7.30% explanation to the variance; it has high loadings on 2 variables. These are family size and total time spent by the household in collecting water. This factor is an index of demography. The higher the number of people in the household, the longer the time spent in collecting water by these households.

vi. Component VI has a contribution of 7.20% to the total variance, with high loadings on uses of water and drying water sources. It is *an index of water availability*. Drying water sources will affect water availability and uses of water. This is because the quantity of available water will determine the uses of water. This particular component stands as an evidence of climate change in Goronyo area.

vii. Component VII has significant loadings on 3 variables, namely: regularity of water source, water borne diseases, volume of water collected, and people awareness of climate change. It is tagged water regularity. This factor has a contribution of 6.60%, it is described as an *index of water reliability***.** The source of water in Goronyo local government area **is** relatively regular, because it is mainly from deep wells and deep hand dug wells. Water in most of these wells last till march, however, by May many of the wells would have dried up due to lowering water table caused by excessive drawdown by over exploitation, evapotranspiration, lack of recharge in the long dry season.

viii. Component VIII is tagged water storage with loadings on 2 variables which are water storage and volume of water collected. This factor has 4.90% contribution. It is tagged *index of water storage***.** The storage capacity of individual households will no doubt determine the amount of water collected.

Table 1: Principal component analysis of water use, climate change and adaptation in Goronyo

	Variables	Components									
		I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**
1	Number Per Household	-.22	-.12	.26	-.15	.73	.07	.00	.05	.28	.02
2	Occupation	-.16	.21	-.06	.09	.08	.12	.16	-.09	.81	.07
3	Number Of Children Under 7 Years	-.27	.06	.02	.31	.62	.04	.03	.01	-.04	-.28
4	Sources Of Water	.05	.10	.12	-.88	-.09	-.12	-.11	.12	-.21	.04
5	Distances To Water Points	.50	.02	.29	.70	-.19	.15	.18	.03	.15	.07
6	Average Time Spent Collecting Water	-.18	-.21	.08	-.06	-.65	-.04	.26	.13	.18	.04
7	Ownership Of Water Source	-.60	.08	-.40	-04	.01	.40	.21	-.20	-.40	.20
8	Maintenance Of Water Source	-.70	.08	-.34	.16	-.05	.50	-.01	.02	-.15	-.01
9	Regularity Of Water Supply	.10	.28	-.06	.14	.16	-.07	.84	-.04	.05	.20
10	Sufficiency Of Water Supply	.50	.64	.02	.12	.22	.15	.20	.15	.04	-.05
11	Accessibility To Water	.43	.80	-.04	.10	.34	.16	-.03	.09	-.01	-.04
12	Water Borne Diseases	.11	-.07	-.41	.12	.23	.01	.70	.12	.26	-.02
13	Payment For Water	-.30	.20	-.02	.80	.11	-.23	.00	.04	-.23	.12
14	Uses Of Water	-.20	.01	.07	.13	.21	.84	-.12	-.07	-.04	.22
15	Climate Change Perception	.09	.24	.31	.02	.32	-.02	-.74	-.15	.02	.05
16	Water Source Drying Up	.21	.05	.21	-.11	.06	-.86	.06	.17	.13	-.13
17	Increasing Dusty Wind	.20	-.10	.09	.10	-.15	.18	.02	-.23	-.24	.80
18	Increasing Rain Storms	.01	-.03	-.12	-.00	-.03	-.06	.10	.21	.19	.81
19	Larger Water Storage	.20	.23	-.09	-.07	-.06	.09	.13	.80	.01	-.08
20	Beast Of Burden	.18	.11	.70	-.21	-.35	-.03	.03	-.23	-.09	-.10
21	Water Conservation Habits	-.12	.19	.80	.19	.12	.08	.10	-.07	-.05	.27
22	Reduction In Water Use	.03	-.08	.80	-.04	.20	.15	.12	.18	-.00	-.21
23	Unclassified Mode Of Responses	.44	.30	-.14	-.05	-.30	-.27	.13	.02	.40	-.17

24	Total Household Time	.93	.13	-.04	-.04	-.10	.07	.03	.10	-.13	.11
25	Queue In Water Points	.93	-.01	-.11	-.04	-.10	.08	.04	.10	-.16	.12
26	Water Collected/Day	-.02	-.30	.16	-.09	.05	.08	-.03	.70	-.40	.16
27	Water Wash Diseases	.16	-.38	-.06	.16	-.05	-.11	.30	.51	.34	.03
28	Conflicts At Water Points	.10	-.60	-.09	.35	.21	.11	-.22	.06	.04	.06
29	Reduction In Agricultural Activities	.06	-.82	.00	-.04	.16	.18	.20	.08	-.23	.07
30	Cottage Industries	.05	-.81	-.08	-.04	-.01	-..10	.01	.05	-.06	-.02

Component Description		Index of Waiting Time	Index of Water access	Index of Water Conservation	Index of Water sourcing	Index of Demography	Index of water availability	Index of water Reliability	Index of water storage	Index of household occupation	Index of Climate change
		IWTIM	**IWACC**	**IWCON**	**IWSOC**	**IHSIZ**	**IWAVL**	**IWREL**	**IWSTO**	**HOCCP**	**ICCAD**
a	**Eigen Value**	**4.58**	**3.62**	**3.00**	**2.74**	**2.18**	**2.16**	**1.97**	**1.47**	**1.28**	**1.08**
b	**Percentage Variance Explained**	**15.2**	**12.1**	**9.70**	**9.13**	**7.28**	**7.20**	**6.56**	**4.90**	**4.27**	**3.61**
c	**Cumulative percentage Variance Explained**	**15.2**	**27.3**	**37.0**	**46.1**	**53.4**	**60.6**	**67.1**	**72.1**	**76.3**	**80.0**

ix. Factor IX is an *index of household occupation.* This factor has a significant loading; household occupation types. It has a contribution of 4.28% to the total explanation. The people occupation has a direct relationship with individual water use habit. For example, occupation is predetermined by the nature and types of education, skill acquisition, access to information, civilization, etc. all these surrogates will affect level of hygiene and the amount of water use.

x. Factor X has the least contribution to the variance (3.61%) in the explanation. Two variables (increasing dust and increasing frequency of rainstorms) have significant loadings on this factor. This factor could be regarded as *index of climate change impact.* Indeed, a cursory look at the study area showed that these are among the dominant impacts of climate change on Goronyo people. The increasing rainstorms of the past 2 years have affected agriculture remarkably in this area. The annual rainfall of 2010 caused large scale flooding in the LGA where some settlements including Sabon Gari Dole were completely washed away, and many acres of onion, garlic, pepper, tobacco, etc were washed away. A condition which has affected food in security.

Relationship between climate change and water use characteristics

The results of the factor regression analyses are presented in Table 2 and 3. These show that the various components offered 84% to the discussion of climate change adaptation and water use in Goronyo. This relationship can be explained further by equation 1

$$y = 2.955 + 0.029IWTIM + 0.0721WACC - 0.092IWCON$$
$$+0.007IWCON + 0.007ISOU + 0.097IHSIZ - 0.0081WAVL$$
$$+0.0071WREL - 0.045IWSTO + 0.005HOCCP + 0.016ICCAD.....eq.1$$
$$R2 = 84.0\%; SE = 0.021$$

Table 2: Component- multiple regression equation of climate change adaption components in Goronyo

Components	Coefficients	T	Standard Error	Percentage Explanation (%)
Constant	2.955	141.8		
1.IWTIM(index of waiting time)	.029	1.355		
2. IWACC(Index of water accessibility)	.072	3.399		
3. IWCON(Index of water conservation)	-.092	-4.386		
4. IWSOU(index of water sourcing)	.007	.338		
5. IHSIZ(index of household size)	.097	4.586		
6. IWAVL(index of water availability)	-.008	-.361		
7. IWREL(Index of water reliability)	.007	-10.598	0.021	84
8. IWSTO(index of water storage)	-.045	-2.151		
9. HOCCP(index of household occupation)	.005	.257		
10. ICCAD(Index of climate change and climate change adaption)	.016	.771		

Component- stepwise regression was also used to rewrite equation 1 with the intention of arriving at the best fitted model. The result is presented in Table 3 and equation 2.

This shows water-source reliability as the most important component contributing 55% to the explanation, this followed by size of household which is large determinant of water use in the area (10%). Others include water conservation methods such as reduction in water use and the use of beast of burden to carry water (9.60 %).

Water accessibility is also important which comprises limited uses of water in both agriculture and cottage industries (5.70%). And lastly water storage component which is relevant to climate change as families now uses larger containers (2.20%)to store water for their uses.

Table 3: Factor- stepwise regression equation of climate change adaption components in Goronyo

Variables	Coefficients	t	Standrd Error	Cum. variance	% Variance	% Total Variance
Constant	2.955	146.17	0.020			
1. IWREL (Index of water reliability)	-.223	-10.93		55.0	55.0	
2. IHSIZ (index of household size)	.097	4.729		65.2	10.2	82.5
3. IWCON(Index of water conservation)	-.092	-4.523		74.6	9.60	
4. IWACC (Index of water accessibility)	.072	3.51		80.3	5.70	
5. IWSTO (Index of water storage)	-.045	-2.218		82.5	2.20	

The above can be modeled by equation 2

$$y = 2.955 - 0.223IWREL + 0.0971HSIZ - 0.021WCON + 0.0721WACC - 0.0451WSTO.....eq.2$$

$$R2 = 82.5\%; SE = 0.021$$

Hence, collection and analyses of data using appropriate procedure is key to the understanding of the different environmental phenomena which are central to man's explanations within his environment.

CONCLUSION

Collecting, analyzing and interpreting data is important in planning and national development. A country and business without data collection will be backward or may be short lived. Hence, data is important. However, if a nation is to develop; effort should starts at the level of awareness on the importance of data collection. It should be stressed that the future of any business outfit and nation lies in collection, organization and interpretation of data.

Explanations of environmental phenomenon such as climate change involves the generation, collection and analyses of large volume of data. Hence, issues such as climate change adaptation involves the use of many variables, in most cases such variables may be closely associated in nature. Multicollinearity must be removed from large sized data before meaningful interpretation will be achieved. The analogy in this paper has quite demonstrated the import of collection, analysis and interpretation of large size data typical in social, environmental and earth sciences.

REFERENCES

Abumere, S.I. (2000) Data analysis, acquisition and analysis. *Paper Presented At The 44th Annual Conference Of Nigerian Geographical Association*, University Of Ibadan.

Adamu, S.O. and Johnson, T.L. (1975) *Statistics for Beginners*. First Edition. Onibonoje Press and Book Industries Nigeria Limited, Ibadan, Nigeria.

Casley, D., and Lury, D. (1987). *Data Collection in Developing Countries*. Oxford, United Kingdom: Clarendon Press.

Chambers, R. (1994) The origins and practice of participatory rural appraisal *World Development* 22 (7): 953-969.

Cochran, W.G. (1997) *Sampling Techniques*. New York. John Wiley.

Groot, A. and M. Maarleveld (2000) Demystifying facilitation in participatory development *Gatekeeper Series* No. SA89, Sustainable Agriculture and Rural Livelihoods Programme

Harvey, D. (1969) *Explanations in Geography*. Sage Publication.

Olorunfemi, J.F. (2004) Methods of data collection in Saliu, H.A. and Oyebanji (eds.) *Research Proposal and Report Writing*. Faculty of Business and Social Science, University of Ilorin.

Selltiz, C., Lawrence, S., Wrightsman and Cook, S.W (1976) *Research Methods In Social Relations*. 3rd Edition, Rinchart and Winston. Simpson. New York.

Slim, H. and J. Mitchell (1992) The application of RAP and RRA techniques in emergency relief programmes, in Scrimshaw and Gleason (Eds.), *RAP Rapid Assessment Procedures* (Boston)INDFC, pp. 251-256.

Tobler, W. (1970) A computer movie simulating urban growth in the Detroit region, Economic *Geography* 26 (2) 234-240.

Tobler, W. (1973) Regional analysis: times series extended to two dimensions. *Geographical Polinica* 25:103-106.