

Smart Digital Census

Authored by: Team of Sapio Analytics & UPICON (Uttar Pradesh Industrial Consultants Limited) led by Shivangi Singhee, Tuisha Agarwal, Ishan Gupta and Aditi Singhai

This white paper introduces the concept of a smart digital census, a form of large-scale survey that provides reports in real time and improves in terms of efficiency and accuracy during its conduction, effectively leading to a cost effective, highly efficient, significantly more accurate, and citizen-friendly replacement of traditional census. Such a revolutionary form of census is presented based on experiences of actual survey done in the state of Uttar Pradesh in India (proof of concept), combined with a detailed deep dive into the history and purpose of census.

Date: 07/06/2021

Contributed by



(Estd. In 1974 By Govt. of U.P. and Financial Institution)

Powered by



Inspired by



Smart Digital Census

A step towards Digital India



Table of Contents

EXECUTIVE SUMMARY	1
1. INTRODUCTION	3
1.1. WHAT IS CENSUS?	3
1.2. ESSENTIAL FEATURES	5
1.3. SIGNIFICANCE OF CENSUS	6
1.3.1. CENSUS – AN EVOLVING TOOL	10
2. CURRENT PRACTICES OF CONDUCTING CENSUS IN INDIA	13
3. GLOBAL METHODS OF CONDUCTING CENSUS	19
3.1. GLOBAL INNOVATIONS	20
3.2. GHANA IS BETTING ON A SMART DIGITAL CENSUS IN 2021 TO TRANSFORM ITS INFORMAL ECONOMY	24
3.3. MULTI PHASE CONTACT STRATEGIES USED IN DIFFERENT COUNTRIES	26
3.4. TRENDS IN THE USE OF TECHNOLOGY	29
4. STATISTICS.....	31
5. ADVANTAGES OF TRADITIONAL CENSUS.....	34
6. LIMITATIONS OF TRADITIONAL APPROACH	35
7. CENSUS TO GO SMART	40
7.1 WHAT IS DIGITAL CENSUS?	41
7.1.1. INNOVATIVE FEATURES OF DIGITAL CENSUS	44
7.2. ADVANTAGES OF DIGITAL CENSUS	49
7.3. DISADVANTAGES OF DIGITAL CENSUS	52
7.4. HENCE, SMART DIGITAL CENSUS	54
8. EVALUATION OF ALL PAPER RESPONSE AND DIGITAL RESPONSE	58
9. CASE STUDY OF A SMART DIGITAL SURVEY IN THE STATE OF UTTAR PRADESH	65
10. BENEFITS OF SMART DIGITAL CENSUS.....	73
11. WHY IS INDIA READY FOR CHANGE?	75
12. CONCLUSION	79
13. REFERENCES	81

List of Tables & Figures

Table 1: Multi phase contact strategies used in different countries	26
Table 2: Trends in the use of technology.....	29
Table 3: Comparison between different types of Census.....	42
Table 4: Evaluation of All-Paper Response & Digital Response	58
Table 5: Complexity of All-Paper Response	59
Table 6: Complexity of Digital Response.....	62
Table 7: On-Ground Survey for Uttar Pradesh.....	72
Figure 1: Census Word Cloud.....	3
Figure 2: Topics on which Census provide information	3
Figure 3: Essential Features of Census.....	5
Figure 4: Utility of Census	6
Figure 5: India's Sex Ratio (Source: Census 2011 Report).....	7
Figure 6: Newspaper article about Finance Commission decision to disburse tax revenue according to Census 2011 (Source: The New Indian Express).....	8
Figure 7: Newspaper article about the decision Of Niti Aayog to decide food subsidy on basis of census data (Source: The Hindu)	8
Figure 8: Article about how census would be the useful tool for marketers in deciding their strategies (Souce: The Economic Times)	9
Figure 9: Census of Beggars in Bombay (TOI Archives).....	10
Figure 10: The Timeline of Indian Census	12
Figure 11: Phases of Census Operations.....	15
Figure 12: Stages in preparatory phase	15
Figure 13: Address Canvassing in US with the help of GIS Maps (Source: US Census Bureau).....	21
Figure 14: Fred Larya interviewing the citizens of Ghana (Source: Ghana Statistical Services).....	24
Figure 15: Main census methods for 2020 round population and housing censuses, percentage	31
Figure 16: Use of Internet in 2010 and Plans for 2020	32
Figure 17: Lessons learned from 2010 round (% of countries by obstacles faced in use of new technologies during the 2010 census round)	32
Figure 18: Lessons learned from the 2010 round, benefits and risks of using new technology in census operations (%).....	33
Figure 19: Transitioning from paper to electronic data collection	33
Figure 20: Census enumerators interviewing the citizens of Mumbai during Census 2011.....	34
Figure 21: Limitations of traditional census.....	35
Figure 22 Cost of census (Source: India Today 2010)	36
Figure 23: Main difficulties in conducting population housing censuses because of impact of the Covid-19 pandemic	39
Figure 24: Innovative Features of Digital Census.....	44
Figure 25: Advantages of Digital Census.....	49
Figure 26: Disadvantages of Digital Census	52
Figure 27: Complexity of all paper response (Source: US Census Bureau)	61
Figure 28: Complexity of digital response (Source: US Census Bureau)	62
Figure 29: Advantages of Smart Digital Survey.....	68
Figure 30: Sample screenshot of real time dashboard	70
Figure 31: Sample screenshot of the application.....	71

EXECUTIVE SUMMARY

Census is an integral process in the growth of any country, as it helps the decision makers and all related stakeholders understand the composition of the nation across demographics, occupation, socio-economic status, and much more. This systematic process of planning, calculating and acquiring information of citizens and then presenting and using them in a manner that can be used for their development is considered critical for all strategies and planning in any nation.

There are multiple types of census created across India, in form of various surveys done by various states as well as the central government spread across multiple parameters, and the national census conducted every 10 years called the Census of India is one of the most important citizen-connect exercises conducted by the Government of India. Data from the Census of India forms the basis of dozens of policies and thousands of decisions taken over the next 10 years.

Being such a critical exercise, Census of India sees a number of enhancements and improvements every decade. Other census activities conducted by different bodies across India have also started seeing a number of innovations to make the process much more efficient and effective, while also making it more accurate as per the requirements of the specific census activity. Traditionally, a pen and paper method where enumerators go to the households of all citizens, with house listing, marking and enumeration being different phases, is the mode of conducting census.

However, a new world order is getting created now. And it's imperative that we listen to these changes happening around us, and ride on the same to bring in the required development in our country.

One of the most critical changes is the digital revolution that is making technology integral to every activity that we are doing now. Another important change is also induced by the pandemic, where we have realised that many activities of the future will need to be conducted while maintaining social distancing and adoption of technologies that can help in the same shall be preferred. We are also seeing more dynamically changing demographics around us, making it very tough to do predictions based on decade old data points. Reverse migration activities during the pandemic pushed these changes, establishing some permanent changes in the

structural distribution of urban and rural India. This is an example which shows that our future will be much more unstructured and will require regular data collection. Policies will need to be made in real time based on fresh data and may need to be modified regularly based on new data. Decisions will not be a function of simple extrapolation of old data anymore, for them to be successful and effective. Such changes are mentioned here in the context of India, but our studies show that this is happening across the world, particularly in the developing economies.

With so many changes, traditional methods of census need a overhaul. This white paper introduces the world to a new concept of “**Smart Digital Census**” which is, as the name suggests, a smarter version of digital census conducted in some parts of the world. The basis of this white paper is a case study from the state of Uttar Pradesh in India, where India’s first smart digital census, and that of the world too based on our global research, was conducted in an attempt to find the problems of the unorganized sector. Conducted across 4 districts of the state, it went to prove the powers and significance of a smart digital census.

In addition to the case study, this paper deep dives into the concept of census and why the concept needs to constantly evolve based on the surrounding changes in any nation. It tries to understand the benefits of a traditional census, and the challenges of a digital census, before introducing the concept of a smart digital census, that tries to learn from the benefits as well as challenges. It defines Smart Digital Census as a well-structured process with multiple steps involving new age technologies (such as artificial intelligence, data analytics, cloud computing, data visualisation, GIS tools, and smart mobile applications), conventional digital technologies (mobile applications, administrator dashboards, and web-based systems) and traditional on-ground enumeration activities; to deliver a deep understanding and analysis of the subjects under census in real time with high efficiency and accuracy due to the self-improving nature of the process.

It also talks about why India is ready to enter into this new era of smart digital census, becoming a pioneer for the world to follow. Through the introduction of this concept, while showing India’s existing advancements in this domain through a case study, it implores the world to list to adapt a concept pioneered in India, while requesting India to adopt this concept across all types of censuses and surveys and see significant financial and social benefits getting generated out of the same.

1. INTRODUCTION

1.1.WHAT IS CENSUS?

Human capital is the most critical capital for contemporary societies' wellbeing and progress. Providing an accurate and reliable assessment of this capital at small area, regional and national levels is of paramount value for evidence-based governing, civil societies, academics and researchers, as well as for all members.

The essential purpose of the population and housing census is to provide that assessment. Aside from the answer to the question “How many are we?” there is also a need to provide an answer to “Who are we?” in terms of age, sex, education, occupation, economic activity and other crucial characteristics, as well as to “Where do we live?” in terms of housing, access to water, availability of essential facilities, and access to the Internet.

The answers to these questions provide a numerical profile of a nation which is the sine qua non of evidence-based decision-making at all levels.

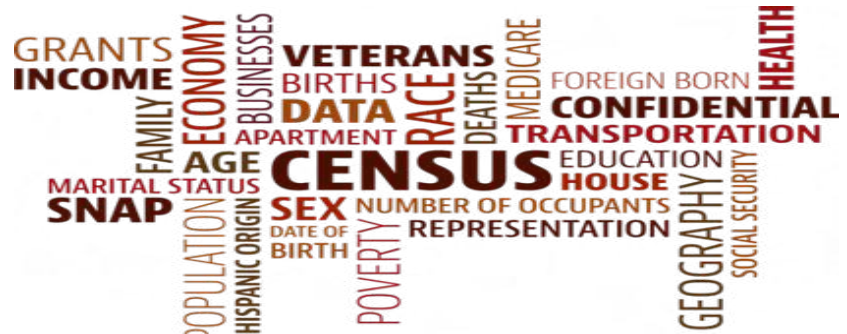


Figure 1: Census Word Cloud

Census is the basis for

- reviewing the country's progress in the past decade;
- monitoring the ongoing schemes of the government and plan for the future;

Census provides detailed and authentic information as explained in Figure 2.



Figure 2: Topics on which Census provide information

Population Census

A population census is the total process of planning, collecting, compiling, evaluating, disseminating, and analysing demographic, economic and social data at the smallest geographical level pertaining, at a specified time, to all persons in a country or in a well-delimited part of a country.

Population is basic to the production and distribution of material wealth. In order to plan for, and implement, economic and social development, administrative activity or scientific research, it is necessary to have reliable and detailed data on the size, distribution and composition of population. The population census is a primary source of these basic benchmark statistics, covering not only the settled population but also homeless persons and nomadic groups. Data from population censuses should allow presentation and analysis in terms of statistics on persons and households and for a wide variety of geographical units, ranging from the country as a whole to individual small localities or city blocks.

Housing Census

A housing census is the total process of planning, collecting, compiling, evaluating, disseminating, and analysing statistical data relating to the number and condition of housing units and facilities as available to the households pertaining, at a specified time, to all living quarters and occupants thereof in a country or in a well-delimited part of a country.

The census must provide information on the supply of housing units together with information on the structural characteristics and facilities that have a bearing upon the maintenance of privacy and health and the development of normal family living conditions.

1.2.ESSENTIAL FEATURES

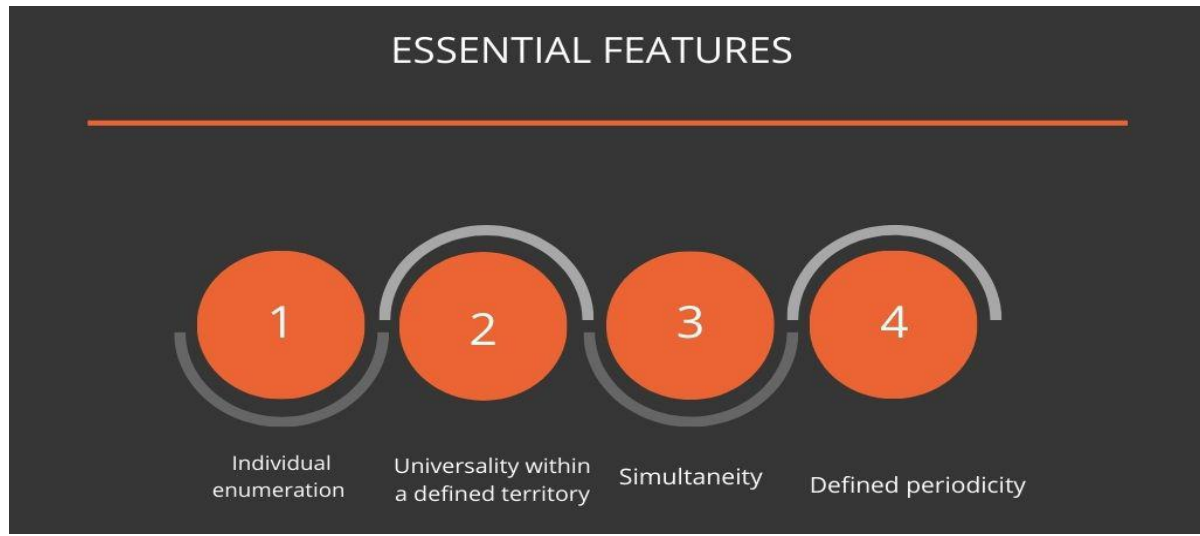


Figure 3: Essential Features of Census

1. **Individual enumeration** -The term "census" implies that each individual and each set of living quarters is enumerated separately and that the characteristics thereof are separately recorded. Only by this procedure can the data on the various characteristics be cross classified. The requirement of individual enumeration can be met by the collection of information in the field, using information contained in an appropriate administrative register or set of registers, or by a combination of these methods.
2. **Universality within a defined territory** - The census should cover a precisely defined territory (for example, the entire country or a well delimited part of it). The population census should include every person present and/or residing within its scope, depending upon the type of population count required. The housing census should include every set of living quarters irrespective of type. This does not preclude the use of sampling techniques for obtaining data on specified characteristics, provided that the sample design is consistent with the size of the areas for which the data are to be tabulated and the degree of detail in the cross-tabulations to be made.
3. **Simultaneity** - Each person and each set of living quarters should be enumerated as of the same well-defined point in time and the data collected should refer to a well-defined reference period. The time-reference period need not, however, be identical for all of the data collected. For most of the data, it will be the day of the census; in some instances, it may be a period prior to the census.

4. **Defined periodicity** - Censuses should be taken at regular intervals so that comparable information is made available in a fixed sequence. A series of censuses makes it possible to appraise the past, accurately describe the present and estimate the future. It is recommended that a national census be taken at least every 10 years. Some countries may find it necessary to carry out censuses more frequently because of the rapidity of major changes in their population and/or its housing circumstances.

1.3. SIGNIFICANCE OF CENSUS

As part of an integrated programme of data collection, the population census is the primary source of basic national population data required for administrative purposes and for many aspects of economic and social research and planning. The value of the census results is increased if they can be employed together with the results of other investigations, as in the use of the census data as a base of benchmark for current statistics.

One of India's great national leaders, the late Shri Govind Ballabh Pant said of the census as follows:

"In fact, in these days you cannot take up any serious administrative, economic, or social work without referring to the Census Report, which is an essential part of every enquiry, of every study. Even for solution of minor problems you have often to consult the Census Reports."

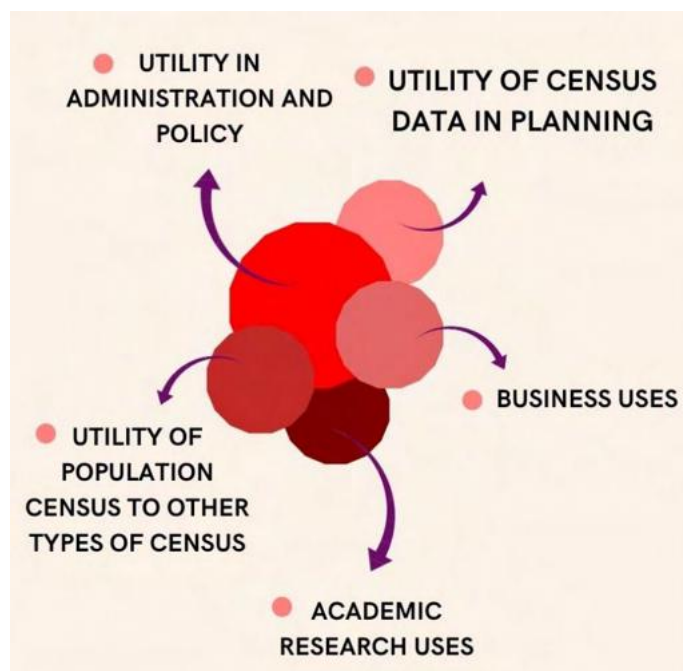


Figure 4: Utility of Census

1. Utility in Administration and Policy

The population census provides the basic data for administrative purposes. One of the most basic of the administrative uses of census data is in the demarcation of constituencies and the allocation of representation on governing bodies. Detailed information on the geographic distribution of the population is indispensable for this purpose. The legal and administrative status of a territorial division also depends upon the size of its population. The social and cultural data collected in the census is employed to determine the total number of seats to be reserved for members of scheduled castes and scheduled tribes in the House of People and the Legislative Assemblies of the States. This information is of immense help for the administration at the district level. Information on the geographic distribution of the population, its size and its other characteristics is essential to the study and evaluation of economic and social problems, which must precede the determination of policy affecting economic and social development.

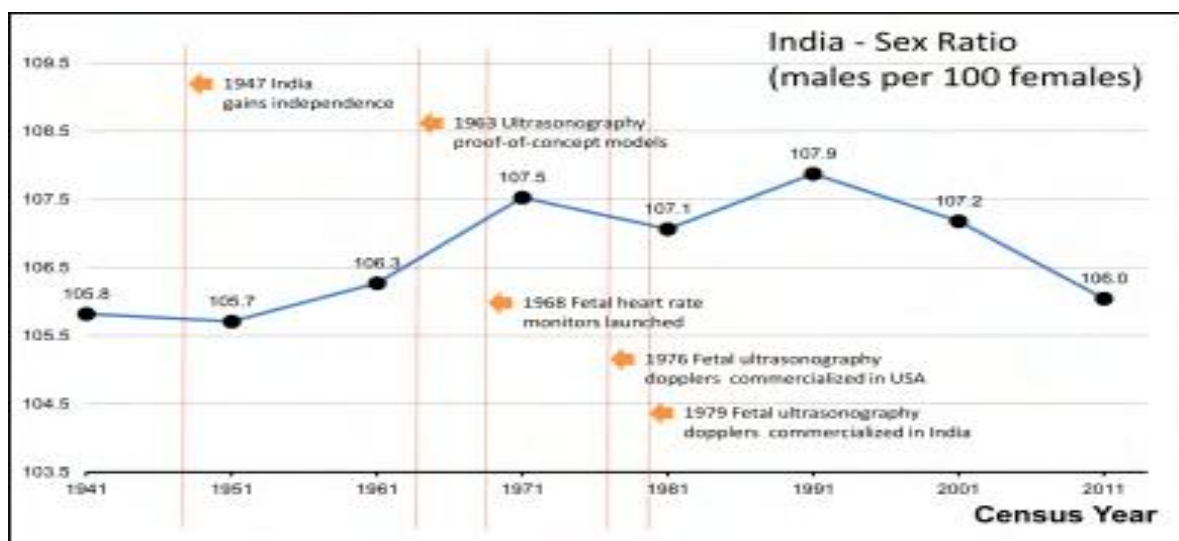


Figure 5: India's Sex Ratio (Source: Census 2011 Report)

The graph is depicting that how the various aspects of country are changing and the same will aid the administration in the policy making exercise.

2. Utility of Census data in Planning

The census data is indispensable for social and economic planning of the Country. **The Planning Commission utilises the Census data on the distribution of population by age, sex classified by rural and urban regions, cities, town areas and social groups to analyse the growth of consumer demand and savings in the process of development.** The census data also prove useful in national income estimates and estimates on differential personal incomes in rural and urban areas and the composition of rural and urban consumption of groups of goods and services and income elasticity co-efficient. **An analysis of areas of different population size with different characteristics certainly serves as a basis for Government plans and investigations in basic social capital.** Besides all these, the census data can prove very useful in the formulation of policies on education, health, agriculture, food and development of road, rail transport etc. In a nutshell, it can be said that the census data is extremely useful for all types of planning



Figure 6: Newspaper article about Finance Commission decision to disburse tax revenue according to Census 2011
(Source: The New Indian Express)



Figure 7: Newspaper article about the decision Of Niti Aayog to decide food subsidy on basis of census data
(Source: The Hindu)

The above news bytes tell how crucial census is when it comes to disbursing of tax revenues or decisions regarding food subsidies

3. Business Uses

Product development and marketing. Accurate census data on where people of different ages live help businesses of all kinds to develop and market their products. For example: Manufacturers of baby products such as baby food, clothes, diapers, and toys, and manufacturers of maternity clothes and greeting cards need accurate information on children as they develop and market their product lines, as do television producers of children's programs and planners of large family amusement parks. Accurate census information on language spoken at home helps television and radio stations define foreign language service areas and marketing companies develop products and services tailored to those who speak languages other than English.

The census data has enough to excite scores of marketers looking to penetrate the many bottom-of-the-pyramid opportunities that may lurk in the hinterland

Thomas Puliyeel, president of market research agency IMRB International, thinks the collection of additional data in Census 2011 will be of significant help to marketing, social and market researchers, planners and administrators. "Data on water availability and sanitation facilities will be of help to those marketing washing products, toilet soaps as well as civic and health planners," (source- economic times(2010))



Figure 8: Article about how census would be the useful tool for marketers in deciding their strategies
(Source: The Economic Times)

4. Academic Research Uses

Accurate census data are vital to researchers in a wide variety of endeavours. Some of the most important needs include the following Any research requiring comprehensive information at the neighbourhood level must rely on the census, with its tract level information. Research on

ethnicity or other research requiring reasonable numbers of observations of relatively small population groups must rely on the census. The decennial census is the only consistent source of data for researchers examining trends over periods of decades. The census is the best source of information for research on immigration and mobility.

5. Utility of Population Census to other types of Censuses

Certain information collected as part of a population census, or incidental to it, can be most useful in conducting and/or utilizing the results of housing, agricultural or establishment censuses taken at about the same time or near about as the population census.

1.3.1. CENSUS – AN EVOLVING TOOL

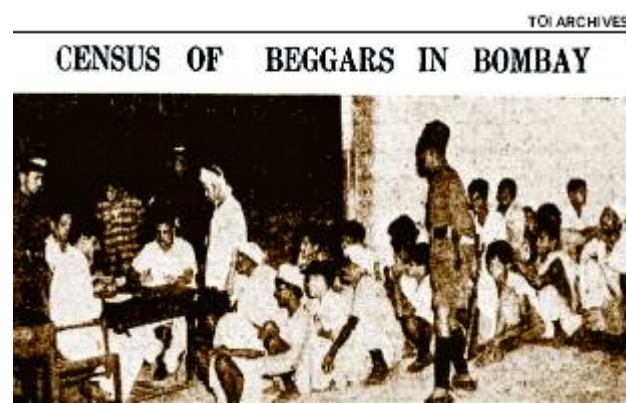


Figure 9: Census of Beggars in Bombay (TOI Archives)

The Census Before Independence

Regional enumerations had previously happened in every administrative area except the Bengal Province. This started with the 1865 Census of the North-West Provinces, which is sometimes referred as the first proper census of India. However, these exercises were usually fragmented

and unsystematic.

The 1871 Census was the first attempt to measure, for all of India, the age, caste, religion, occupation, education and infirmities of each person. However, Berar, Oude and the Punjab were purposely left out. Officials thought it was unnecessary to spend the money in those areas, since they had been measured recently through regional censuses.

The censuses were designed more for social engineering and to further the British agenda for governance rather than to uncover the underlying structure of the population.

The sociologist Michael Mann says that the census exercise was "more telling of the administrative needs of the British than of the social reality for the people of British India." The difference of the nature of Indian society during the British Raj from the value system and the societies of the West were highlighted by the inclusion of "caste", "religion", "profession" and "age" in the data to be collected, as the collection and analysis of this information had a considerable impact on the structure and political overtones of Indian society.

The creation of the Census was driven mainly by the British emphasis on administration, rather than a genuine desire to uncover the characteristics of India's population. In fact, the first Census in 1871 was partially a reaction to the 1857 Rebellion. The rebellion had posed a serious challenge to British rule, so the Raj prioritized administration and governance as a means to stability after 1857.

The next eight censuses were characterised by mistrust and confusion. Various rumours spread through the Indian population — that the censuses were meant to help the British introduce new taxes, convert the population to Christianity, create a register for military or labour recruitment, or document forced migration. Violence occurred, though this was more common in places where tensions were already high. Some enumerators even recounted being attacked by tigers.

The Census After Independence

After Independence, the census was guaranteed through the Census of India Act 1948, which lays down the rules and regulations on conducting the national census. The Act also makes it obligatory for the public to answer all census questions faithfully, while guaranteeing the confidentiality of those answers.

The 1951 Census was the first census held under an independent India, and the ninth overall. It recorded India's population as 361 million people, with 83% of the population living in rural areas.

Since Independence, the census data has been used by many people and organizations for many purposes. The government uses the data for administration, planning, making policies, and managing and evaluating various public programmes. It is also used to demarcate constituencies and allocate representative to Parliament, state legislative assemblies, and local bodies.

At the same time, non-government actors have found value in the census data. Researchers and demographers use the data to predict population growth and find demographic trends. Businesses and industries use the data to plan their penetration into underutilized areas.

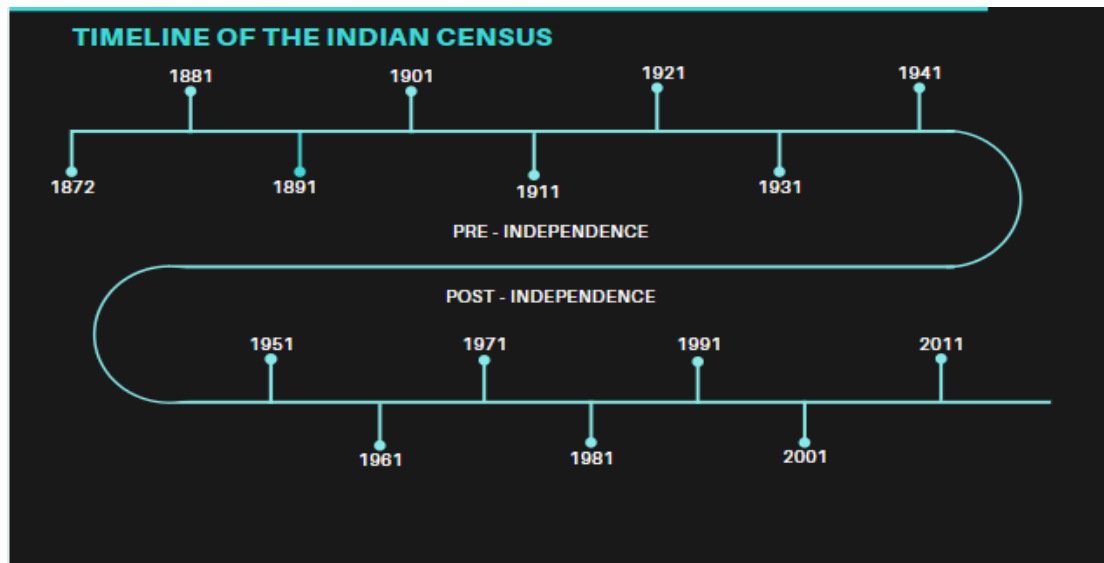


Figure 10: The Timeline of Indian Census

This timeline shows the various years in which the census was conducted in pre-independence era and post-independence era. The census has also played a prominent role in identifying social and economic trends for almost a century and a half. For instance, the first census in 1872 highlighted how low literacy and social attitudes limited the ability to recall age, as most people preferred to round up the number either to 0 or to 5.

The Census Act came into force in 1948 (Act No.37 of 1948). The censuses of post-independence era were conducted as per the provisions of this Act.

2. CURRENT PRACTICES OF CONDUCTING CENSUS IN INDIA

Census Division is responsible for conduct of Census Operations once in ten years under the legal authority of the Census Act, 1948 and the Census (Amendment) Act, 1993.

The Census Operations in India have been carried out in two phases-

1. House listing and Housing Census:

In House listing and Housing Census phase, buildings and structures, residential, partly residential or non-residential are identified and listed systematically. Housing Census in India provides statistics on housing stock, availability of basic amenities in each household and condition of human settlements. This information helps in assessment of housing deficit, besides overall quality of dwellings and people living therein. This helps to prepare a complete and unambiguous frame for the Population Enumeration.

Following are the parameters included in the House listing Operations:

- Use of the census houses
- Condition of census houses used as residence
- Predominant material of the roof, wall and floor of the census houses
- Type of structure of census houses
- Number of dwelling rooms
- Ownership status of the house
- Number of married couples and whether they have independent sleeping rooms
- Source of drinking water (e.g., Tap; Hand pump; Tube well; Well; Tank; Pond River; canal; Spring; Other) and its location
- Source of lighting (e.g., Electricity; kerosene; Solar energy; Other oil, Any other; No lighting)
- Availability of bathroom, type of latrine and type of drainage for wastewater
- Availability of separate kitchen and type of fuel used for cooking (e.g., Firewood; Crop residue; Cow dung cake; Coal, Lignite, Charcoal; kerosene; LPG; Electricity; Biogas; Other)

- Availing of banking services and availability of the specified assets (e.g., Radio, Transistor; Television; Bicycle; Motorcycle, Moped; Car, Jeep, Van; None of these.

2. Population Enumeration:

The Population Enumeration follows the Housing Census within a gap of six to eight months which counts each person in the country, Indian national or otherwise, along with her/his individual particulars and information on social, cultural, demographic, economic and other aspects.

Following are the parameters included in the Population Enumeration (For each member of the household):

- Relation to head of the household
- Sex
- Age at last birthday
- Current marital status
- Age at marriage
- Religion
- Name of Scheduled Caste / Scheduled Tribe
- Mother tongue, other languages known
- Highest educational level attained
- Disability status
- Worker / non –worker
- Main Worker / Marginal Worker
- Place of Birth
- Reason for migration

The population census is one of the most extensive and complicated statistical operations which needs careful planning to avoid defects and inefficiencies. Careful planning of census is vital to the successful operation of the census. The census operations can broadly be divided into the following sequential phases:

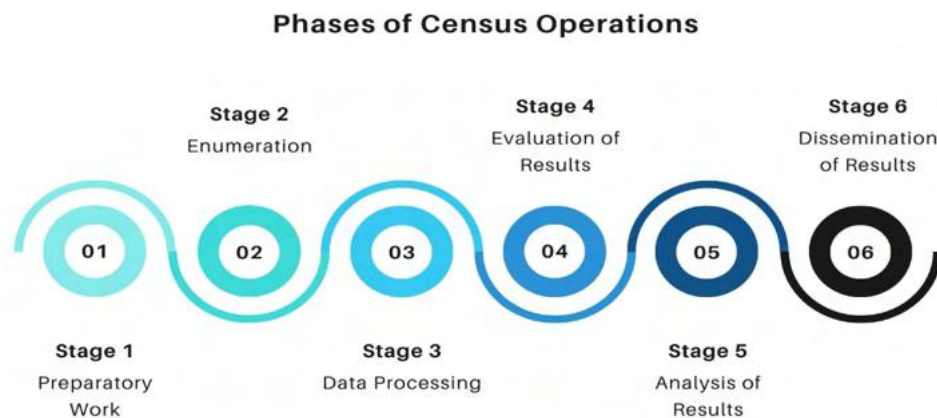


Figure 11: Phases of Census Operations

1. Preparatory work:

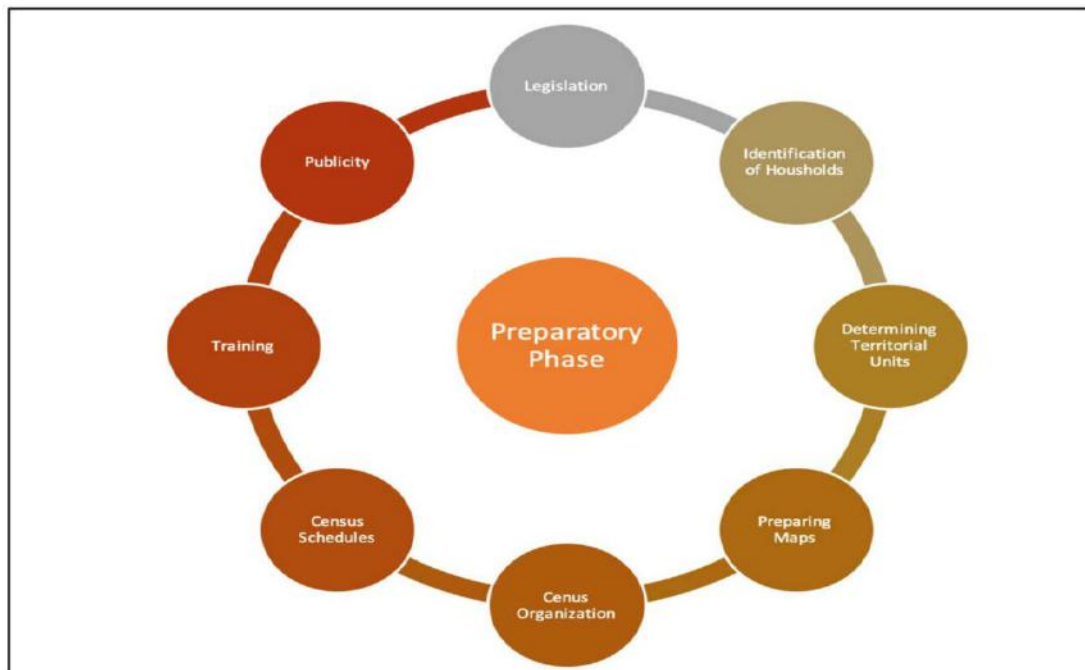


Figure 12: Stages in preparatory phase

The preparatory work is the first and most important step, which includes:

- **Enactment or modification of Census Legislation:** The Census Organization should be armed with necessary authority to have access to households and canvas the prescribed questionnaires and to expect the people to answer truthfully. The law should also protect the interests of the people by guaranteeing the secrecy of the information

collected. In India - a permanent legislation, the Census of India Act was placed in the Statue Book in 1948.

- **Creation of administrative organization:** Government to notify the date for the census and to appoint a Census Commissioner and Superintendents of Census Operations in States.
- **Demarcation of administrative units:** A fundamental task in preparing for a census involves the location of every local area and building up of a suitable organization to adequately cover it, without omission or overlapping.
- **Household, the Operational Unit of enumeration:** The operational unit is the physical entity of the household which is generally understood as a group of persons commonly living together. According to the Census of India Act, it takes obligatory on every person occupying a house, enclosure etc. to allow access to census officers, and to allow them to paint on or affix to the place such letters, marks or numbers as may be necessary for census purposes.
- **Formulation and pretesting of census questionnaires:** Early in the preparatory stages of the census, action would have been taken to evolve the census questionnaire after duly consulting the various departments of the government and the other census data users and to pretest them and finalize them. The pretests should have helped to fix the norms for coverage by each enumerator within the enumeration period.
- **Staff recruitment and Training:** Thorough training of census enumeration agency from top to bottom, each one in the census hierarchy must be very clear on what his duties and responsibilities are. There should be no ambiguity in their understanding of census concepts. Manuals of instructions issued by the Chief Census Commissioner or the Regional Commissioners are of be of great help.
- **Publicity:** Wide publicity programs are launched to educate the public on the need for the census, about the particulars that will be asked at the census, assuring the people that the confidentiality of the information will be maintained and also to tell them of their legal obligations. All media of publicity are used such as issue of attractive pamphlets and striking posters, radio and television programs, by feature films and cinema slides, through a short lesson in schools, by announcing by beat of drum in villages and various other media according to the ingenuity of the local officers.

2. Enumeration:

The Census Enumeration process is of mainly two types:

- Canvasser method the enumerator approaches every household and records the answer on the schedules himself after ascertaining the from the head of the household or other knowledgeable persons in the household.
- Householder method the enumerator distributes the census schedules to each household in his jurisdiction and the head of the household is expected to fill the answer for all members of his household and the enumerator later collects back the answered schedules soon after the census day is over.

It has to be recognized that in countries where literacy is still low, the Canvasser method is the only practical method.

- **Sampling in Enumeration:** Sampling may be used for the collection, tabulation and evaluation of census data. Sampling, therefore, plays an important role in population censuses both as an integral part of plan execution, analysis and evaluation of census data and through the use of census as a sampling frame for subsequent sample enquiries.
- **De- jure or De-facto Enumeration:** An important question pertaining to enumeration that always arises at a Population Census is whether the population should be counted on a de-facto basis i.e., at the place where a person is actually found on the reference date of the census or on a de-jure basis i.e., count a person only according to the place of normal residence. Generally, in practice, enumeration on a 100 per cent de-facto or de-jure basis is impossible and often a variation or even a combination of the two is resorted to. The census instructions should clearly lay down who are the persons who should be enumerated during the census enumeration period.
- **Special Cases:** Nonsynchronous enumeration, snow covered areas, Person on Boat, person on long travel, houseless people.

3. Data Processing:

The processing of the census data occupies a very important place in the census. The census enumeration may be thorough and accurate, but the census tabulation will not be accurate and useful unless the raw data are properly processed. There are various methods of data processing

and the choice of a suitable method depends upon the circumstances of the country. A complete census is a huge undertaking and will require many years to complete the whole tabulation program of the census data. By the time the data is published, it becomes out of date and cannot be used for any policy matter.

4. Evaluation and Analysis of Results:

The evaluation of census results in respect of completeness and accuracy occupies an important place in the census program. The extent of error can be estimated using checks of the internal consistency of the data, by examination of the reasonableness of the results and by comparison of the results with data collected in other enquiries. The publication of the census results must specify the extent of error in coverage and reporting. It is the responsibility of the census organization to analyze the census data so tabulated. This will greatly help to make comparative studies and indicate the long-term trends of certain characteristics of the population. The experts from outside the census organization should also be invited or associated with the projects dealing with the analysis of census data.

5. Dissemination of Results:

The census would be considered incomplete if the data tabulated in the desired form is not available to the data users in the form in which it suits them. This calls for high priority to be given to the publication of the census data. The Census Organization should make every effort to publish important results as early as possible so that their usefulness and the extent of their interest may not be diminished. It is very necessary to fix the target dates for publication well in advance so that there may not be any difficulty in achieving them. Along with publishing of the Census data, it would be proper that every country may provide a methodological and administrative report.

6. Systematic Recording of Census Experience:

The experience of the past census team always proves useful in the planning of the new census. As the census is conducted every ten years, there is the possibility that the experienced staff may leave the census organization. For this reason, it is very essential to assemble complete records on the methodology of each census, an evaluation of the technique employed, and the costs.

3. GLOBAL METHODS OF CONDUCTING CENSUS

MOMENTS AND MILESTONES IN GLOBAL CENSUS TAKING

The modern population census began to evolve only in the 17th century. Before that time, inventories of people, taxpayers, or valuables were certainly made, but the methods and purposes of such inventories were different from modern ones. The most important difference was that early inventories were made to control particular individuals—e.g., to identify who should be taxed, inducted into military service, or forced to work. Because it was usually not to an individual's interest to be counted or to give correct information for these purposes, the premodern enumerations tended to be inaccurate. A second difference was that early inventories did not seek to count all the people or even a representative sample of them but only those in particular categories, such as family heads or males of military age. Such surveys are known to have been made in ancient Babylonia, Palestine, Persia, China, and Egypt. Every five years, the Romans enumerated citizens and their property to determine their liabilities. This practice was extended to include the entire Roman Empire in 5 BCE. After the collapse of Rome, the practice was discontinued in the West until the modern period.

The modern idea of a population census as a complete enumeration of all the people and their important characteristics for purposes of understanding the basic structure and trends of the society rather than for identifying and controlling individuals slowly arose in the 17th and 18th centuries. There is no such thing as “the first census” because, although early efforts embodied one or another modern feature, none of them embodied all modern features. Perhaps the first effort in an area larger than a city to count everyone at successive intervals was made in New France (Quebec) and Acadia (Nova Scotia), where 16 enumerations were made between 1665 and 1754. Only gradually did the early census takers learn what was basic information and how to obtain it. The first United States census did not, for example, secure data on occupation, birthplace, marital status, or exact age.

The history of modern census taking thus involves **three parallel developments**:

- (1) The invention and gradual spread of the idea of countrywide enumerations for general scientific and governmental purposes
- (2) The improvement of administrative machinery, techniques, and accuracy of enumeration, including legal safeguards assuring that individual answers would be kept in confidence
- (3) The deepening and systematization of the types of information obtained. All three developments tended to increase the effectiveness of censuses as sources of essential information.

3.1. GLOBAL INNOVATIONS

Innovation 1: Using Satellite Imagery to Check Addresses

“The world we live in is increasingly challenged. Changes in population, over-consumption, sustainability-these are some of the things we can begin to understand with GIS and data. The 2020 Census gives us the opportunity to collect better and more accurate data that can help us with understanding. With the 2020 Census, many more countries, including the United States, are using GIS to bring together maps with rich data sources combined with powerful analytic tools to bring better understanding.”

President American Census Bureau

Innovation 2: Introducing Online Self-Response

Innovation 3: Introducing Mobile Devices for Enumeration

INNOVATIONS USED FOR 2020 DIGITAL CENSUS (USA)

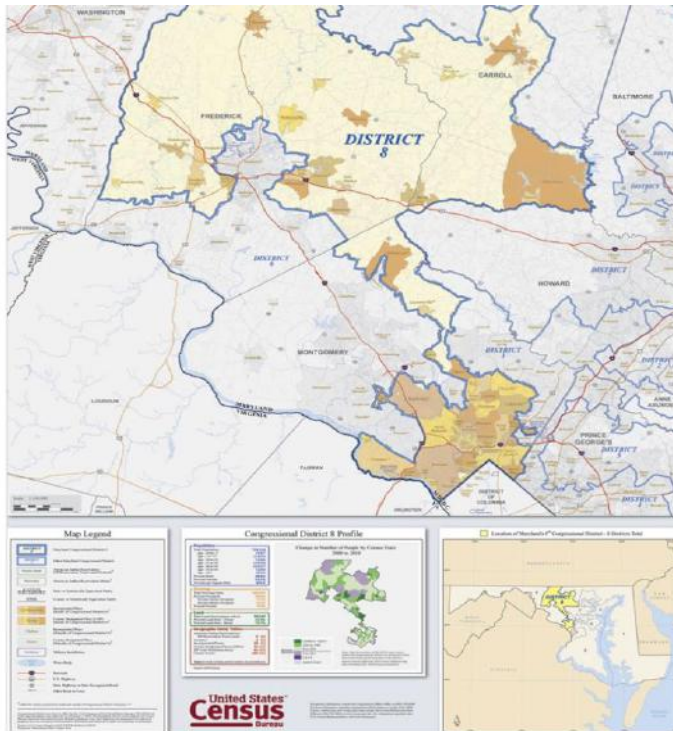
Innovation 1: Using Satellite Imagery to Check Addresses

Before the Census Bureau can count every person in the country, it must first collect addresses for every housing unit. One way the Census Bureau uses this address list is to mail census materials, including invitations to respond online, by phone or by mail. Census Bureau employees used to “canvass” neighborhoods in person, jotting down new addresses and correcting old ones on paper.

This long-running operation, known as Address Canvassing, is one of the ways the Census Bureau updates its Master Address File or MAF. Address canvassing was costly and time-consuming. Employees travelled a total of 137 million miles to update the MAF before the

2010 Census. **In 2015, the Census Bureau began using aerial images from a network of satellites.**

The Census Bureau developed computer software that allows employees in offices to compare satellite images from 2010 to new ones taken in real time. This helps them identify new houses, apartment buildings and other units to verify in the traditional Address Canvassing operation.



The US Census Bureau's increased use of GIS will help states redistrict their jurisdictions quickly.

Figure 13: Address Canvassing in US with the help of GIS Maps
(Source: US Census Bureau)

validated 65% in the office, removing them from the in-field workload. That means workers needed to canvas fewer neighborhoods in person, saving time and money.

Innovation 2: Introducing Online Self-Response

The 2020 Census was the first time everyone had the option to respond to the census online as well as by phone or mail.

The Census Bureau had an Internet Self-Response tool designed to make it easy to complete the questionnaire online and keep responses secure. Directions for responding online would be included in letters, postcards and other mailings sent to most homes beginning in mid-March.

Every response submitted on the internet shall be encrypted. That means data are changed into a code that only Census Bureau data analysts can read. **Responses travel through a secure cloud computer network and the Census Bureau locks them in a “digital vault”.**

Innovation 3: Introducing Mobile Devices to Enumeration

From collecting census responses and job applications to storing questionnaires, the Census Bureau used millions of pieces of paper to gather and file information. It relies much more on technology – and much less on paper. **In 2020, census takers who go door-to-door to help people respond would collect information on smartphones using a custom application created by the Census Bureau. To protect privacy, administration encrypt all data and devices require two-factor authentication to be unlocked.**

When a device connects to the internet, encrypted data immediately transmits to the Census Bureau’s digital vault – and is no longer on the device. Encrypted data are only stored on the devices until they connect to the internet. Software in the smartphones also provides specific routes for census takers to follow to visit homes. Optimizing routes in this way helps census takers do their jobs more efficiently. If a device is lost or stolen, the Census Bureau will remotely wipe it clean of all applications and information.

Ethiopia embraced Chinese technologies during their Fourth National Population and Housing Census, which was scheduled to take place from April 7 to 28 2019.

The Ethiopian Central Statistics Agency (CSA) finalized to undertake Ethiopia's first ever digital technology-assisted national population and housing census. The CSA had previously revealed the purchase of close to 180,000 tablets from Chinese technology giants, Huawei and Lenovo, to effectively conduct the population and housing census, which is held about every 10 year.

The products, which had been tested and approved by the agency, were circulated across Ethiopia as the East African country embarked on its first-ever digital population and housing census procedure, which includes the use of mobile tablets for data collection through five major local Ethiopian languages.

Experts at the CSA had previously told Xinhua that the use of mobile tablet technologies during the data collection procedure *"will immensely reduce various difficulties that may affect the census process."* The new technology is said to effectively establish careful data registration system, which will eventually avoid duplication and double counting during the data collection and registration procedure, according to experts at the CSA. The census will be conducted using five local languages namely Amharic, Afaan Oromo, Afar, Tigrigna, and Somali.

The application of digital mobile technology is also said to be a positive impetus to incorporate different aspects of communities, including the use of five languages for data registration purpose unlike the previous proceeding.

3.2. GHANA IS BETTING ON A SMART DIGITAL CENSUS IN 2021 TO TRANSFORM ITS INFORMAL ECONOMY

A CASE STUDY

Fred Laryea, a grocery shop owner in Accra, the capital city of Ghana, goes from house to house with a black backpack. He talks to residents inside in the local language, Twi, earning their trust and asking very detailed questions about individuals living there and amenities available to them.



Figure 14: Fred Laryea interviewing the citizens of Ghana (Source: Ghana Statistical Services)

“How many people live here?” “What’s the date of birth for every one of them?” “Can you read and write?” “Do you use a private toilet?” As Laryea asks each question, *he records the answer on a tablet, and then taps “next” on the touchscreen to go to the next question.*

This is Laryea’s job for the week: he is one of the temporary workers hired by Ghana Statistical Services (GSS) to help carry out a trial census currently underway in a few selected districts across different regions in Ghana.

The trial is to prepare for the next decennial census in 2021. *Equipped with a tablet loaded with a digital neighbourhood map, and a portable charger, field workers like Laryea are testing new technologies moving away from the pencil and paper of the previous census.*

The field workers are saving and sending answers back via an app on the tablet to the central server in real time and verifying where interviewees live by GPS. Araba Forson, the deputy government statistician at GSS, say that using tablets will save the processing time down the road, as well as improve coverage and data quality.

For a country like Ghana, where data on its residents are incomplete and infrequently updated, the census provides a rare opportunity to collect data on the people who are missed in other forms of data collection and map every housing structure in the country. *Aided by digital*

technologies, the census results can be useful to allocate resources and formalize much more of the country's economy

DIGITAL AMBITIONS

The scale of data the census collects is crucial to the government's effort to formalize larger parts of Ghana's economy, where over 80% of the employed non-farming population work in informal sectors, according to the 2015 Ghana Labour survey. Jobs such as street vendors, taxi drivers or head porters, most of whom are female migrant workers from rural regions working in urban areas, aren't registered with the government. There's no effective way of tracking people in Ghana.

The country has tried to roll out a national identification system multiple times in the past without success. There isn't a uniform address system, either: Many in rural areas do not have a numbered address and rely on a nearby landmark such as a tree, or a gate, to indicate where they live. The current administration wants to solve both problems: The government is pushing residents to register biometric national ID cards (Ghana cards) that assign Ghanaians unique ID numbers, as well as building a formal addressing system to document all housing structures in the country since 2017. *By digitizing housing structures and recording household GPS locations, the census will provide data to accelerate both processes.*

A complete database of its people and buildings could help inform policy decisions on taxation and social welfare for the government: a step toward a more formalized economy.

The government plans to spend \$84 million for the operation, according to GSS, a step-up from \$54 million ten years ago. GSS, with a staff of about 300, will hire an estimate of 60,000 temporary field workers for the three weeks of census operations.

3.3. MULTI PHASE CONTACT STRATEGIES USED IN DIFFERENT COUNTRIES

Table 1: Multi phase contact strategies used in different countries

AUSTRALIA (2016)	• PHASE 1: Most households received a letter with a unique Census Login, and instructions on how to complete the Census online.
	• PHASE 2: Reminder letters were sent to every household that hasn't responded to the Census
	• PHASE 3: Visit from Census Field Officers paper forms still available to households that require them and could be requested through a dedicated paper form request line. Each paper form was delivered with a reply-paid envelope.
CANADA (2016)	• WAVE 1: A letter, received eight days before Census Day, inviting respondents to complete the Census by Internet using a personal Secure Access Code, and offering the possibility of requesting a paper questionnaire by calling a toll-free number. • WAVE 2: A reminder letter, received just after Census Day, reminding non-respondents to complete the Census questionnaire (also offering the possibility of requesting a paper questionnaire);
	• WAVE 3: A questionnaire packaged received about one week later. This package contains a new access code for Internet. The letter accompanying the package contains more direct wording concerning the mandatory requirement to complete the census.
	• WAVE 4: Telephone voice broadcast messages sent after Wave 3. Phone or personal visit by a census enumerator start about one week after reception of Wave 3.

ESTONIA (2011)	• WAVE 1: e-census for one month, where all permanent residents of Estonia can complete the census questionnaires online by themselves.
	• PHASE 1: mid-September: sending of letters to dwellings selected in the sample. Thereafter, people may respond to the census questionnaire online.
	• PHASE 2: October: mailing of paper questionnaires to non-responding dwellings • PHASE 3: October – November: sending of reminder letters to non-responding dwellings • PHASE 4: November/December: field enumerators conduct the Building Census (all dwellings).
	• PHASE 5: January/March: field enumerators complete collection of population census information (sample).
UNITED KINGDOM (2011)	A fully traditional 100 per cent field enumeration, self-completion Census
	• PHASE 1: Mail-out of paper questionnaire packs 7-19 March to ensure delivery one week before Census Day. Questionnaire included Internet access code for those respondents wishing to use online option.
	• WAVE 2: interview census where enumerators visit those people who did not complete the questionnaires during the e-census. Interview census lasted about seven weeks.
	There was a pause of 18 days between e-census and interview census for rearranging the working lists of non-enumerated respondents
JAPAN (2015)	• PHASE 1: Enumerators distributed internet IDs to all households

	• PHASE 2: Enumerators distributed questionnaires to households which have not responded through the Internet. Questionnaires were submitted by handing over to enumerators or by postal mail
	• PHASE 3: Non-response follow-up by field enumerators
KOREA (2011)	• PHASE 1: Internet option available for 10 days just before Census Day.
	Incentives included handing out gifts by drawings and giving students two-hour credits for volunteer work.
	• PHASE 2: Field interviews during 15 days just after Census Day.
LATVIA (2011)	• STAGE 1: During 12 days, residents had opportunity to fill in the questionnaires on the Internet, authorising with the identity code and passport number, Internet bank ID of five banks and e-signature. To ensure the safety of personal data, after two days, it was not possible to sign in using the identity code and passport number.
	• STAGE 2: Field interviewers visited residents at their homes. Stage 2 started five days after stage 1 and lasted 2 and a half months.
	• STAGE 3: Considering the high interest of the population and its willingness to fill in the questionnaires on the Internet, as well as approval of the Population and Housing Census commission, a third stage was implemented, in which residents not taking part in the 1st and 2nd stage of the Census had an opportunity to fill in the Census questionnaires on the Internet or to visit some of the 37 regional centres and fill in the questionnaire. Stage 3 took place immediately after stage 2 and lasted 10 days.
LITHUANIA (2011)	• STAGE 1: e-census for 16 days, where residents were afforded an opportunity to enumerate themselves electronically.

- STAGE 2: a population survey – interview, where those who had not participated in the e-Census were visited by interviewers. Stage took place about three weeks after Stage 1 and lasted about one month.

3.4. TRENDS IN THE USE OF TECHNOLOGY

Table 2: Trends in the use of technology

Census Rounds				
	Before 1990	1990	2000	2010
Mapping	• Sketch	• Digital • Sketch	• Digital • Sketch	• Digital • Geo referenced data • Sketch
Enumeration	• Paper questionnaire	• Paper questionnaire	• Electronic questionnaire – Internet • Paper questionnaire	• Electronic questionnaire – Internet, hand-held devices and telephone • Real-time monitoring paper questionnaire
Data Processing	• Manual data capture	• Optical data capture • Manual data capture	• Optical data capture • Manual data capture • Capture	• Optical data capture • Automatic data capture (Internet and hand-held devices) • Web-based manual data capture • Manual data capture

Dissemination	• Paper publication		• Paper publications • Online digital publications	• Online interactive databases • Digital publications • Paper publications
----------------------	---	--	---	--

(Owing to the COVID-19 Pandemic, data for 2020 round was not available as the census is yet to take place)

4. STATISTICS

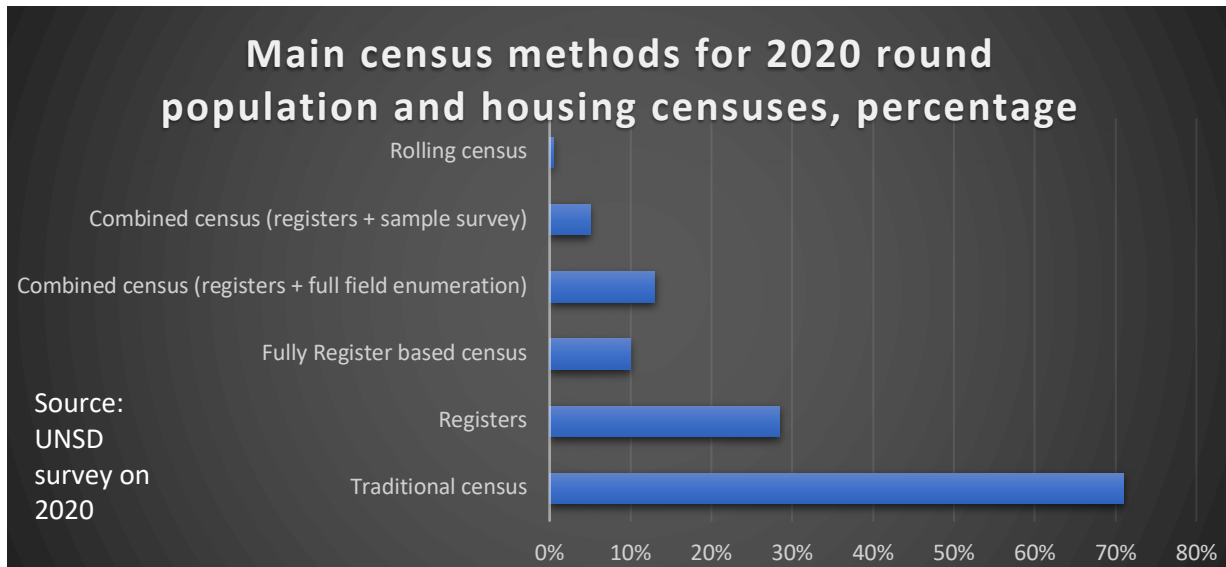


Figure 15: Main census methods for 2020 round population and housing censuses, percentage

The graph above depicts the census methods for 2020 round population and housing censuses (percentage wise). The results indicate that the traditional census, based on full field enumeration whereby each household in the country is approached with a request to provide information, is the predominant methodology in conducting population and housing censuses. Out of **158 countries** that responded to the survey, **112 countries (71 per cent)** used or will use the traditional approach to censuses. A significant number of countries (**45 countries out of 158**) had used or will use administrative registers for producing small area statistics, either using administrative registers with full field enumeration (**13 per cent**) or combining registers with sample surveys (**5 per cent**) for compiling data on census topics that are not available in administrative registers. Among countries which use administrative registers, only **10 per cent fully rely on administrative registers** for producing small areas statistics. **The rolling census methodology is so far used only in one country.**

Figure 16 depicts the use of Internet for census in the year 2010 and plans for 2020(postponed due to the Covid-19 pandemic) across the continents.

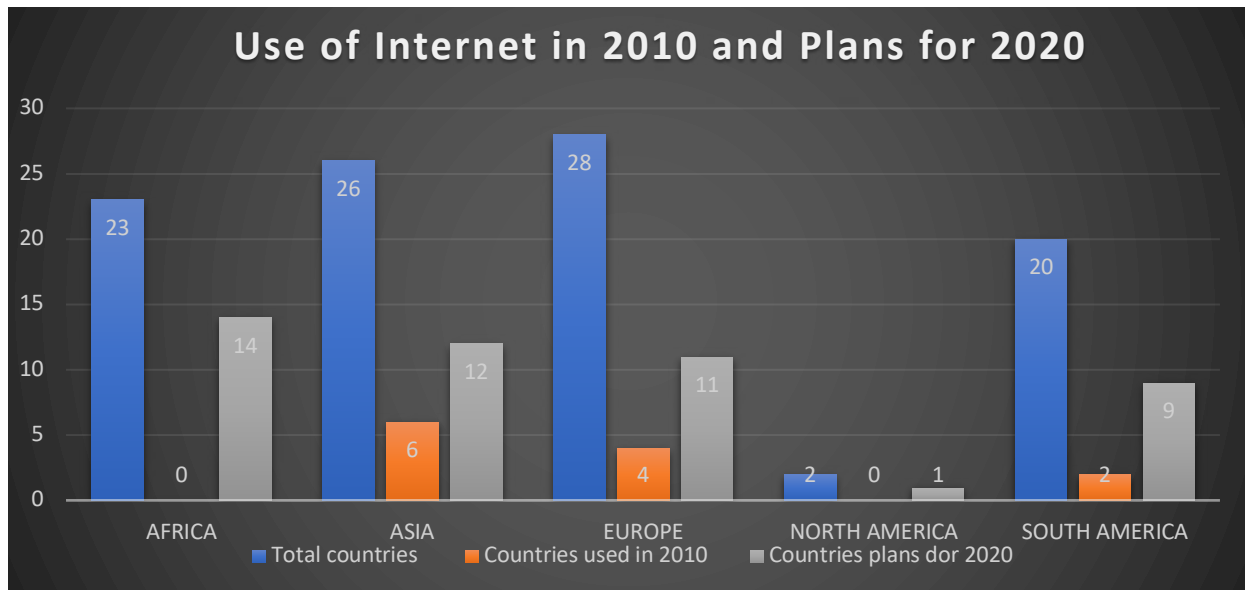


Figure 16: Use of Internet in 2010 and Plans for 2020

While Africa had none of the country in 2010 which used internet to conduct census. For 2020 census, 14 countries of Africa have planned to conduct census with the use of Internet. Asia has 6 countries in 2010 which used internet for census, the number has doubled for census 2020. In Europe and South America there is a stark jump in number of countries which have planned to go smart in 2020 census.

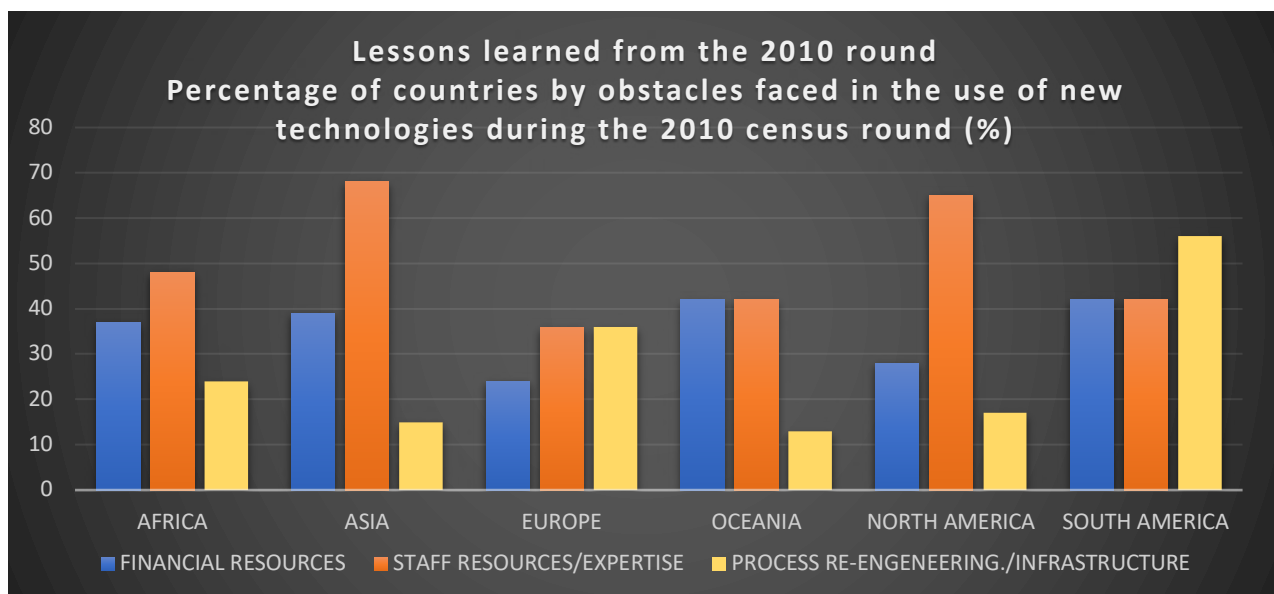


Figure 17: Lessons learned from 2010 round (% of countries by obstacles faced in use of new technologies during the 2010 census round)

This graph shows the obstacles faced by different countries (continents) around the world in using the new technologies for census.

All over the world, the leading obstacle has been to find the right staff resources and expertise while conducting the survey, followed by financial resources and Process Re-engineering varying largely in different geographical areas. Most of the continents faced problems in staff expertise following it was the problem of financial resources.

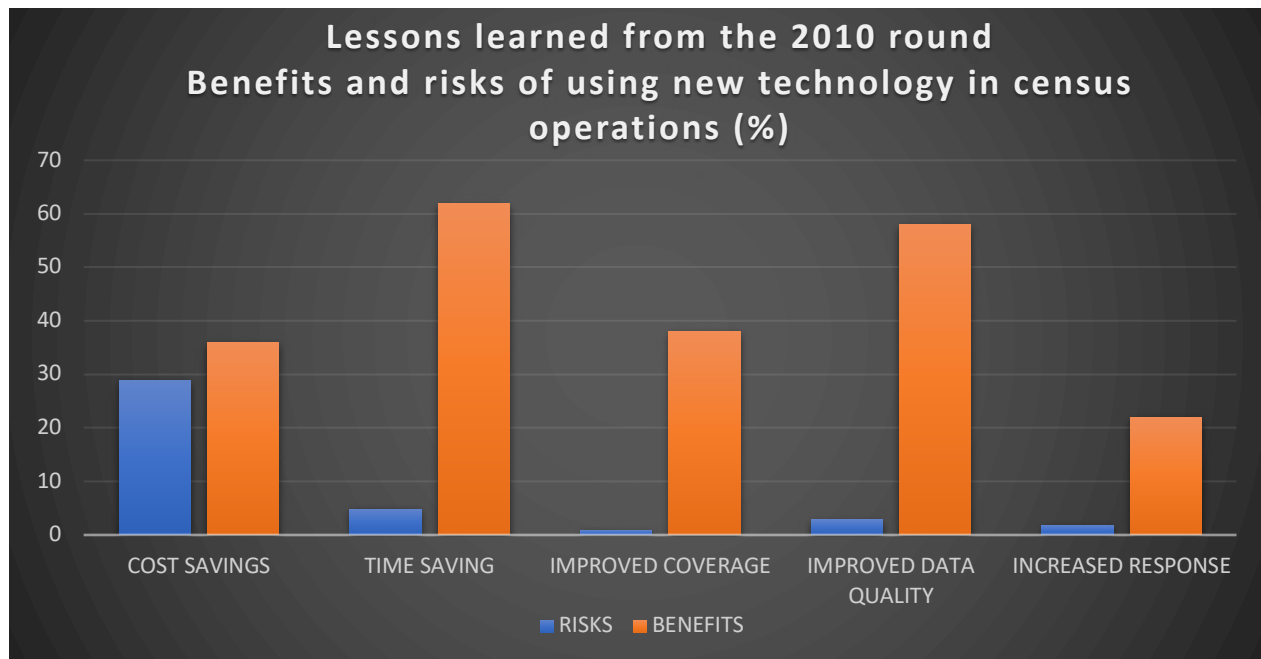


Figure 18: Lessons learned from the 2010 round, benefits and risks of using new technology in census operations (%)

This graph shows a comparative analysis of Risks and Benefits of using new technologies for the purpose of population enumeration. Showing overall more benefits of using advanced technology in terms of Time Saving, Improved coverage, Improved data quality, increased response as well as cost saving.

TRANSITIONING FROM PAPER TO ELECTRONIC DATA COLLECTION

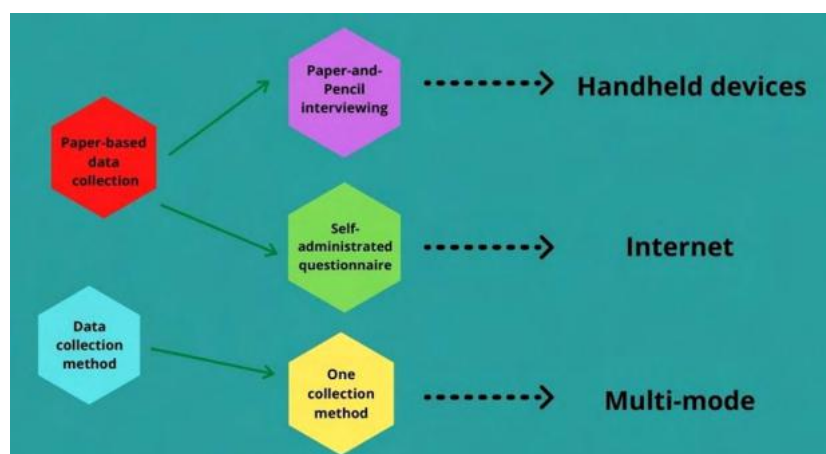


Figure 19: Transitioning from paper to electronic data collection

5. ADVANTAGES OF TRADITIONAL CENSUS



Figure 20: Census enumerators interviewing the citizens of Mumbai during Census 2011
(Source: Census Digital Library, Government Of India)

The **comprehensive nature of the traditional census** in terms of topic coverage should not lightly be overlooked when compared with some of the alternative approaches. And there is the positive spin-off which the census generates for official statistics in general, through its public awareness campaign. As well as engendering a sense of national participation it presents an opportunity to the government through its **spokespersons to highlight the importance of objective and timely statistics for society at large**. Traditional census also **ensures simultaneity**. Respondent answers the questions simultaneously as soon as enumerator puts the question, and the recording of responses is done. Another major advantage is that there is lesser need for complex data adjustment since processing of raw data provides all inputs.

This method also ensures **coverage of population across the length and breadth of the nation**. It is possibly easy to reach out to the vulnerable sections of the society who are not literate or equipped technology. **On field enumerators helps them in the interpretation of questions and thus avoiding the chances of misleading responses**. In addition to this, since it is face to face census, emotional quotient of respondents can be understood while answering the questions. In case of traditional census, it is **perceived that data theft may not be easy since the data is running across crores of pages in the physical format**. This somewhere helps in gaining the confidence of citizens who would otherwise be worried about their credentials being getting leaked and their security being at stake.

6. LIMITATIONS OF TRADITIONAL APPROACH

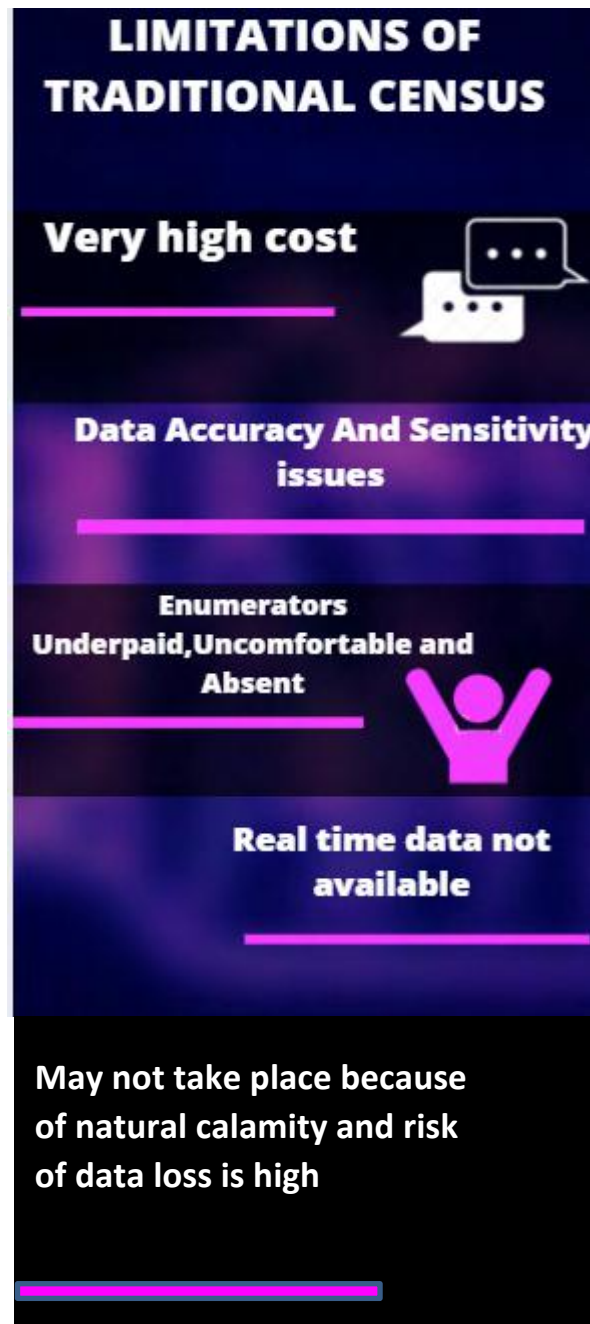


Figure 21: Limitations of traditional census

1. Very high cost:

The census conducted in a traditional way is very expensive. **The main cost item is for the temporary work force (enumerators, supervisors, etc.) that has to be recruited and trained, and has to work for a few weeks or longer periods.** Considering that an enumerator is needed on average for about 100 households (this is a rough estimate, as the real number depends on the characteristics of the census, the territory and other factors), it is clear that the

number of persons to be recruited is very high, and so is the cost. **Apart from the cost for census field staff, the cost of printing, distributing, collecting a huge number of census forms, entering the data (manually or using scanners) and processing them is also very high.**

The whole exercise of Census 2011 will cost approximately Rs 2,200 crore while the NPR will cost Rs 3,539.24 crore. Thus Rs 7,000 crore--along with the Registrar General of India's (RGI) own expenses--will be spent on this exercise. The Union Cabinet last week approved the expenses of the NPR. Considering the normal growth rate of our population, India will be home to over 120 crore people by March 1 next year, the cut-off date for the census. That means roughly Rs 59 will be spent per person for this head and house counting and preparing the NPR. While inaugurating the All India Conference on National Population Register, Chidambaram said, "Nowhere in the world has a government tried to count, identify and issue identity cards to more than a billion people. This is perhaps the biggest exercise of its type since humankind came into existence." Says C. Chandramouli, registrar general and census commissioner, "Our job will be complicated. We are going to knock every door and



"Our job would be complicated. We will knock on every door of the 24 crore households in India and count every head within a given time limit."

C. CHANDRAMOULI, Registrar General & Census Commissioner

Figure 22 Cost of census (Source: India Today 2010)

The above article appeared on India Today news portal which discusses about the cost of census

Census 2011 took place in 3 phases:

1. House listing and Housing Census (Apr 2010- Sept 2010, 6 Months)
2. Data Collection (9 Feb - 28 Feb 2011, 20 Days)
3. Revision Round (1 March - 5 March 2011, 5 Days)

The expenditure on census was Rs 2,200 crore. 733,600 crore tons of paper was used during the Census.

Not only the very high costs, but also **the cost distribution over time and the peak around the period of the fieldwork can create problems for the management of the traditional census.** In many countries, it is **difficult to recruit many temporary census staff for the fieldwork operations, considering that they must have the necessary skills but can be**

employed only for a short period. In many National Statistical Institutes, once the census operations have been completed, it is not possible to retain the staff that worked for the census; they are often reallocated to other services or released. In this case, the knowledge accumulated while planning and conducting the census is lost unless the same staff can be re-employed for the next census.

2. Data Accuracy: Issues Around Sensitive Questions

One major issue raised was that census enumerators did not necessarily ask all relevant questions. People reported that census enumerators filled questions about gender without asking the question. This became an issue for transgender people who might look outwardly like one gender but consider themselves the other gender.(Census 2011,India)

Even when these questions were asked, families could be hesitant to disclose sensitive or embarrassing information. For example, usually the head of the household answered Census questions for all people in the household. He might report someone as his son, even if that “son” considers herself a woman. **The same issue arose around disability data. The Census added more detailed questions about disabilities, and even ran a campaign around accurately reporting disabilities.** Yet, since disability is a taboo subject, households were often inclined to hide disabled family members when Census enumerators showed up. The result was a low, likely inaccurate, percentage of disabled persons in the country.

3. Enumerators: Underpaid, Uncomfortable and Absent

Working at such a massive scale inevitably led to problems because of the number of people involved — **25.4 lakh people worked on the Census, 21 lakhs of whom were enumerators.** Three years after the Census of 2011, many teachers in Nagpur (India) reported that they had not yet received their full salary. Census enumerators were supposed to receive 18,300 Rs. and supervisors 24,300 Rs. each. However, enumerators and supervisors had only received 8,300 and 12,300 Rs. respectively. By 2015, the rural development department of the Maharashtra government had dispersed 5.5 crore rupees for rural enumerators, but it had not yet sanctioned the 2.4 crore rupees required for enumerators and supervisors in Nagpur.

In addition, many enumerators in Mumbai — home to Kamathipura, the largest red-light area in the world — reported discomfort with their work. The enumerators said that having to go to Kamathipura and visit brothels was “embarrassing”.

Another problem that arose was the prevalence of proxy enumerators in Andhra Pradesh, where enumerators outsourced their work to other people. In Hyderabad, the Times of India reported that an enumerator left his forms with a neighborhood's watchman, promising to return in 4 days to collect the filled forms. In Vijayawada, teachers passed their Census forms off to their students, and some officials in Visakhapatnam passed off their forms to colleagues. The problem is that proxy enumerators were not properly trained, so the quality of the data they collected was generally low. While the state's census director made trips to some districts to oversee the exercise, some observers say she turned a blind eye to issues with the enumerators (reference: media reports)

4. Real time data not available

Data visibility is not in real-time, to process complete census is a huge undertaking and will require many years to complete the whole tabulation programme of the census data. By the time the data is published, it becomes out of date and cannot be used for any policy matter.

Data Processing took around 1.5 years using the ICR technology in 2011 census of India. The updated data was published in May 2013 (reference: Census India website and media reports) and various reports have been published as late as 2016, using the census data.

5. Census may not be conducted due to natural calamities

Census activities are generally not possible in the events of natural calamities which put the safety of workers at stake. Due to COVID-19 pandemic many countries of the world postponed their Census taking. Also, there are certain chances of destruction or loss of records due to fire, natural calamity like flood or earthquake.

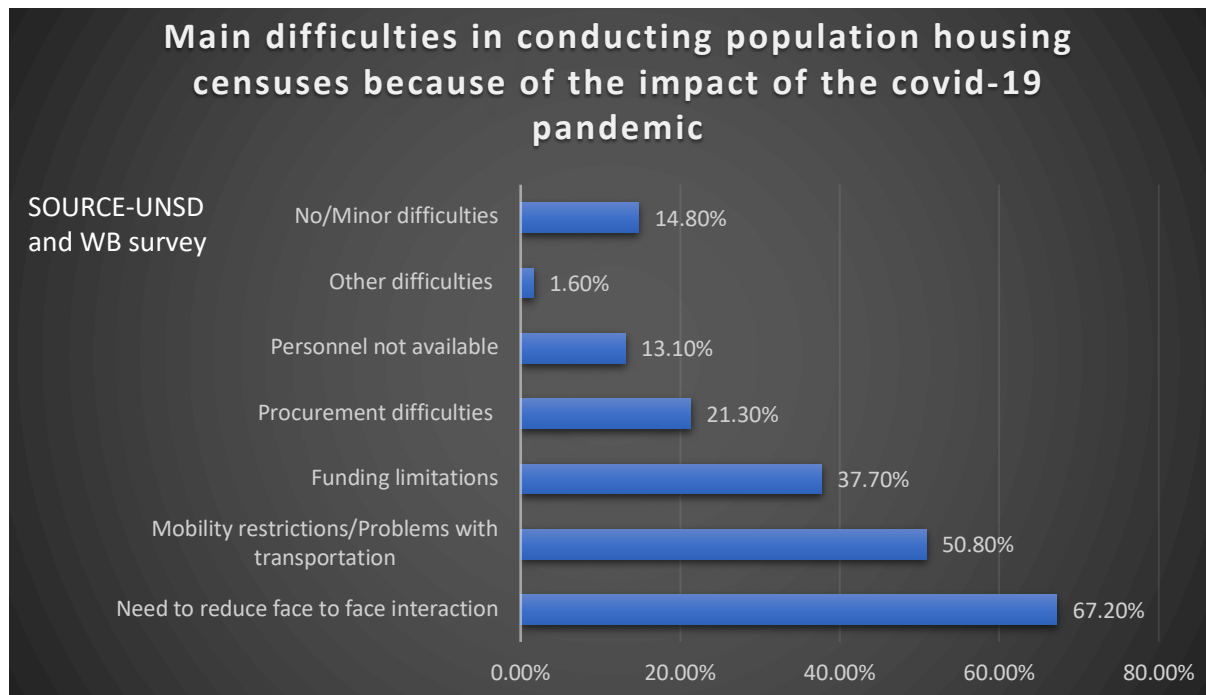


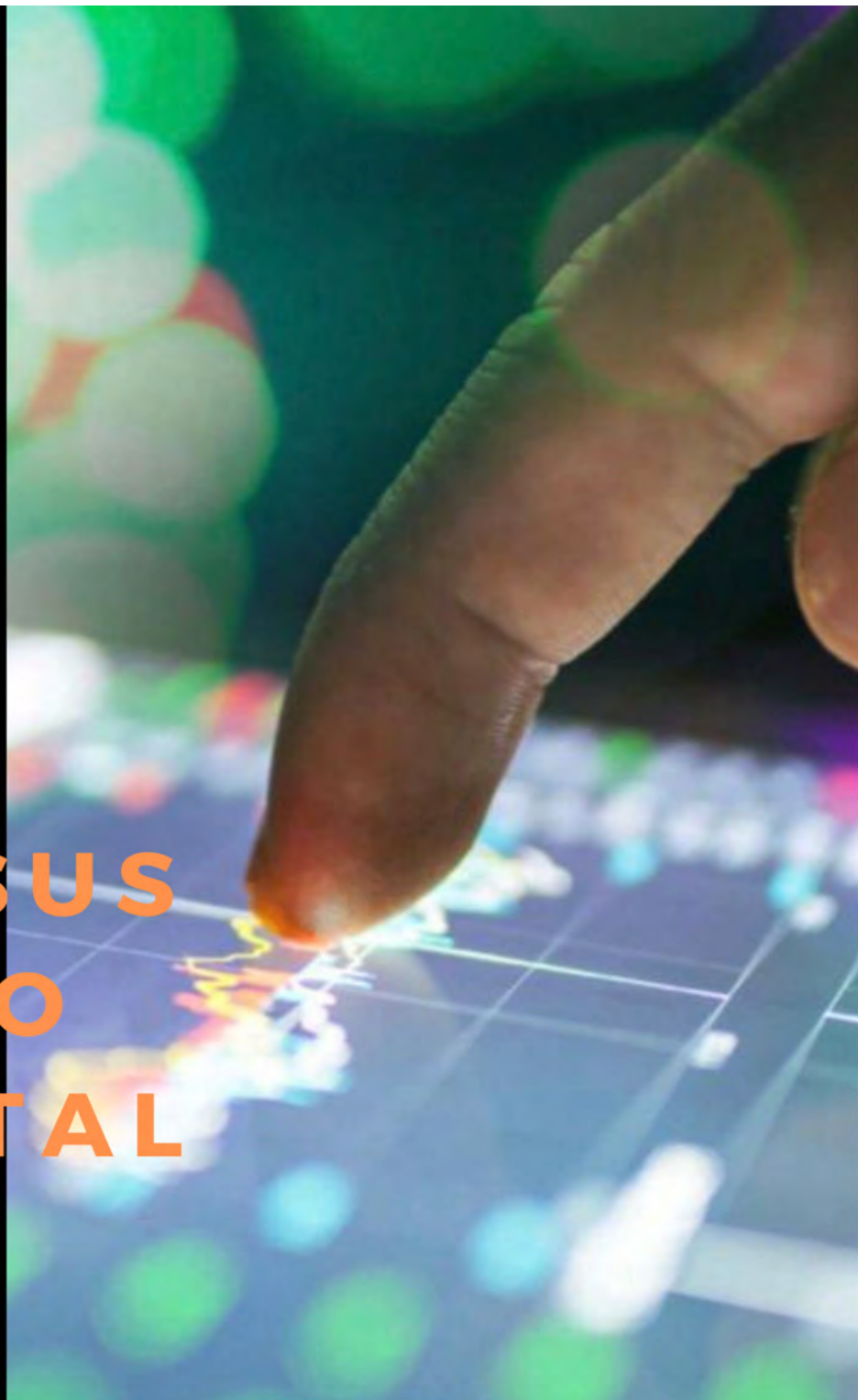
Figure 23: Main difficulties in conducting population housing censuses because of impact of the Covid-19 pandemic

From the analysis of this graph, it can be deduced that face to face interaction, which is the biggest aspect of traditional census, served as the major difficulty in conducting the census of 2020.

6. The risk of respondent data loss is high

1. Responses lost due to mail or sorting problems
2. Handwriting that cannot be interpreted 100% accurately
3. Stray markings that otherwise result in a misread, preventing complete data capture.

CENSUS TO GO DIGITAL



7.1 WHAT IS DIGITAL CENSUS?

It is commonly known that the art of population census taking goes back many centuries. Ever since the end of the nineteenth century, there have been efforts to take advantage of a succession of newly available technologies to make such large and costly statistical enquiries more efficient and effective. A census is labor-intensive, requiring large numbers of temporary staff. Personnel costs usually are the principal component of census budgets, with expenditure for information and communication technology coming second. Even small improvements in the methodologies used, or in the effectiveness of the equipment, can result in important gains in quality and/or cost-effectiveness of the whole operation.

Digital Census is a method of collecting the census data using new technology advancements like through a mobile application, a web-based portal.

Smart Digital Census takes conventional digital census to the next level, by using unique technologies that can help improve the process of census while it is being conducted, and in the process making the entire process smarter and much more effective than a conventional digital census. While a digital census is limited to use of digital technologies while collecting data from the citizens, a smart digital census does much more, by using many more innovations that tie up with the process of data collection.

Think of it this way: what if the enumerator on ground doing data collection using the digital technology is also supported by an artificially intelligent guide who is constantly making the process easier for them; while also making the outputs easier for the administrators in real time.

Table 3: Comparison between different types of Census

	Traditional Census	Digital Census	Smart Digital Census
Pre-work	Regular	Regular	Very High, as systems are customized to become guides for enumerators based on secondary research; also, auditing and other innovative mechanisms need to be defined right at the start of survey
Data Collection	Enumerator goes on ground with pen and paper, collects data, and later the paper data is transferred to a central database	Enumerator goes on ground with a digital system (mobile application or a laptop), collects data, and data immediately syncs with the central database	Enumerator goes on ground with a digital system (preferably, mobile app), gets guided by the app itself to do faster data collection, and data immediately syncs with the central database and an auditing system that improves the process through real time or daily recommendations

Data Outputs	Data normalization and standardization takes up a lot of time, before outputs are created	First level outputs are expected to be created in real time, and then improvements may be required before final submission	Outputs are created in real time (initial checks, auditing, visualization, all happen at the same time)
---------------------	---	--	---

7.1.1. INNOVATIVE FEATURES OF DIGITAL CENSUS



Figure 24: Innovative Features of Digital Census

1. Use of Mobile Apps with Templates

Imagine a Mobile App that has the feature to guide the user on the next steps using templates that can be fed into the application. So, an application that is used by an enumerator to collect unorganised sector in a remote village can have a different template while collecting this data, and the same application can have a different template while collecting data from corporates in a metro city.

Effectively, one need not create multiple applications and systems for different exercises, as the same can be used for different use cases, through intelligent use of templates. It can also be used

to conduct different census surveys at different places with different questionnaires using a same template app.

Some of the benefits that applications driven by templates provide are:

a) No Reinvention for repetitive features across different needs

Several features are repetitive, and there is no need to reinvent the wheel. Here are some examples of features:

- Sign-In
- Sing-Up
- Reset Password
- User Class
- Installation Class
- Role Class

The examples above are only a small portion of features that are repetitive, and you can avoid using an app template.

b) Saves cost

Since the UI already exists in the template, government do not need to assign a sizable amount of money on the development of the frontend. This is one of the most expensive portions of app development, where governments spend lacs of rupees.

Think from the perspective of conducting numerous surveys across multiple departments and sectors in a large country like India, and we will realize how these costs can become significant.

c) Bug-free development

The first draft during development is always fraught with several bugs and malfunctions. Developers then need to go through each of these bugs and determine the cause for the improper functioning before they can eradicate the bugs. This lengthy process takes time and immense effort. In templates most of the coding is already free from bugs or other issues. Therefore,

time can be saved, produce bug-free apps while also lowering the overall budget for the application.

d) Less Time for new creation or for changes

We need to be ready for the world that's changing at a very fast pace. One application may need to undergo numerous changes based on changing situations within the same region. Such changes can be handled through applications that are driven through templates, as no new development shall be needed. The structuring of templatization should be done in a manner that this aspect is taken care of.

2. Real time output, audit and recommendation

a) Prompt formulation of policies

Government Units will not have to wait for another decennial census to take place so that they can decide the certain aspects of policies like the beneficiaries to be part of welfare programmes, the umbrella of people under Below Poverty Line etc. This will be possible since the data is directly pushed into the database, cleaned and visualised on the census dashboards without much human intervention.

For example –A NITI Aayog paper has recommended reduction in National Food Security Act coverage. They are waiting for new census data to be available, so that Centre may consider revising the number of people who get subsidised food grains under the National Food Security Act. If real time data would have been available, it would have made the process faster.

b) A move For Multipurpose Identity Card

"Why can't we have just one card for all utilities like Aadhaar, passport, bank account, driving licence, voter card. There should be a system that all data should be put together in a single card. This is a potential. That is why digital census is very important," Amit Shah, Home Minister of India.

This may also link the registration of birth and death with the country's voter's list. A citizen would not have to apply for a voter card once they reach the age of 18; it would happen on its own. Similarly, the voter list would be updated on its own.

c) Unified Model of Public-Private Partnership

Since the data would be visible on real time basis, it will help the administration to track the changes in the population characteristics and in administering and formulating of policies. Citizens will be updating the information and government would not need to keep crores of files lined up, it will also prevent the loss of data.

3. GIS Maps for Address Canvassing

a) Every head will be counted

During enumeration exercise, enumerators will have the GIS Maps in their devices which will help them to reach out to every settlement, so no one is left uncoded. It will avoid non - inclusion and double inclusion of units that may result in under or overcounting. Also tailored communication and partnership campaign can be designed to motivate people to respond in the remote parts of country as depicted by GIS maps.

b) A great tool for administration

This data once completed will be displayed on maps. Incorporating GIS and spatial technologies will make data more usable and in accessible format. A click will aid the administration to know the whole demography of the population residing in the particular location.

4. Feedback Loops to Increase Output

a) Help govt to have more coordinated, collaborative and committed deliverables

Constant feedback loops will help the government to reorient existing systems as well as shape the new ones. Bottlenecks will be alleviated without the need to replace whole system. This will help in promote SDG goals under the cluster of peaceful and inclusive society.

b) User Experience

It will help the administration to deduce the problems the users are facing while using the app. This will help them to quickly rectify the issues so that the interface can become more and more user friendly. This will increase the rate of responses as well.

5. More Surveys Per Enumerator

There will be pre-determined fields and scroll down option in the app and if the enumerator is taking the response of a particular area, he will not have to fill the fields again and again, they will be auto filled, saving their time. Also, the app has the capability to ask the next question basis the response to the last question i.e. redundant questions can be completely eradicated.

This becomes more effective when additional smartness is applied in form of artificial intelligence, as explained later under smart digital census.

7.2. ADVANTAGES OF DIGITAL CENSUS



Figure 25: Advantages of Digital Census

1. Logistics involved with printing, mailing, managing, archiving and scanning paper form gets eliminated

When these logistics gets eliminated, it leads to saving of the funds and those funds can be channelised towards other development processes. Lesser workforce is needed who would otherwise been involved in performing these tasks. It becomes cumbersome for the offices to keep crores of files safely. With the advent of digital census, there is no need to keep so much of data in the papers.

2. No need to convert from paper to digital

Since the data is already stored in the mobile devices in the digital format, there is no need of paper. This also avoids the cumbersome task of storing the data from paper to digital format, which takes lot of time and reduces the efficiency of workforce. It saves the paper also.

In 2011 Census Of India, 7,33,600 crore tonnes of paper was used. Digital census can turn this number almost to nil, which itself is the great driving force.

3. Integration is seamless and near real time

Power of real data can be harnessed to innovate for the citizen, in partnership with the public sector's ecosystem. Government can become data-driven and make evidence-based, strategic investment and policy decisions, change management needs to view and utilize data as an asset.

Government can extract timely insight when analysing policies and access a wealth of historical and operational data from across government to understand which factors cause a public problem, what types of programs they should invest in to solve the problem, and whether the chosen policy is having the intended impact. Also, administration can monitor and review their performance in real time to see if they're working as designed – or if critical design assumptions have changed, requiring program modification.

Citizens want all their key information available in one place and accessible through a smooth interface. In a data-driven organization, citizens only need to register once to access critical services. With a single version of the truth, that registration can extend across every interaction the citizen has with your organization.

4. Reduces human error through standardisation

Since apps are already templated, possibilities of human error get reduced. All the people must do it, is to choose from scroll down list for certain basic features and enter details in predetermined list. There is not also no barrier of handwriting problems and whether they will be misunderstood during final analysis. This also increases the efficiency of human capital.

5. Provides respondents with instant language and customised help

The apps will be having the various language options, so the language barrier will be eliminated. Also, in case of any problems during enumeration by enumerator or self-enumeration customised help option like screen tips will be present on the app itself which will clear any doubts regarding the questionnaire.

7.3. DISADVANTAGES OF DIGITAL CENSUS



Figure 26: Disadvantages of Digital Census

1. **Not physically tangible like paper**

There is a sense of comfort with paper which has been developing over a long period of time. Since the data is in digital format, it won't be tangible like that of paper.

2. **Fear of confidentiality and security of data**

Confidence in government is crucial to a stable and productive society. As a government works to become more data driven, it must maintain this confidence. However, there are a range of challenges that are encountered on this journey. Below are the key barriers

- Access and use of citizens data and the associated privacy aspects
- Managing the sharing of data within government and across agencies
- Enhancing cybersecurity to create a secure environment for all
- Managing data and data quality to assure confidence in the resulting outcome

In the same way that paper questionnaires may be lost at any point from their completion to their delivery, census information submitted electronically is also vulnerable to risks specific to digital data. Too many online responses at any one time may cause the server hosting the collection portal to crash, and respondent data may be corrupted or lost in the process. Staggering the delivery of questionnaires or invitation letters or requiring people outside predetermined target populations/areas to contact the census agency before they can use the Internet form, may be a means of restricting use of the Internet form at any one time.

There is also a risk that respondent data may be intercepted during transmission or compromised when delivered to the statistical agency. Special measures must be taken to ensure that digital data will be delivered via a secure connection from an external network to an internal storage location, as well as to prevent malicious attacks from the outside.

3. Increased customer service/helpdesk capabilities

During management of information respondents may need to seek the help of trained human personnel. This would require the administration to increase the help desk services for the same. This will add up to the cost as in the salary of workers and the infrastructure required at the service centres.

4. Higher initial cost

The process to develop and test an Internet data collection system and online questionnaires is long and complex, and the initial investment to introduce electronic collection requires significant financial resources and expertise. Net cost savings may not always be achievable, at least initially.

For the first-time implementation, take-up rates may be difficult to estimate since tests and pilot censuses are not always representative. Also, given the rapid advancements in technology,

it is not safe to assume that the same technology can be reused 5-10 years down the road without at least some revision or upgrades. There are, of course, ways to reuse the existing technology, but further investment will usually be necessary to upgrade to new technologies and to expand Internet take-up capacity to meet increasing demand.

5. Lesser reach for older population and low-income households

It is possible that the older population and low-income households may not access to digital world or may not be digitally savvy. So, in such cases there is the possibility of lower response rate.

Government can roll out certain policies in this area after studying the exact characteristics of population and can deploy on field enumerators who can fill the forms in their own tablets and can guide them.

7.4. HENCE, SMART DIGITAL CENSUS

The Smart Digital Census targets all the disadvantages in the conventional digital census and removes most of them.

Smart Digital Census is where enumerators go on ground to take surveys, but they use new age technology enabled systems to do the same, with a lot of technology-driven pre-work that makes their process easier, faster and much more accurate, while also making it real time for immediate actions driven out of recommendations out of the surveys.

Let's take the concept of digital census which is conducted through a mobile application (with templates), real time outputs/auditing/recommendations, GIS mapping, and even feedback loops to constantly monitor the census process.

Now, consider the following additions to this concept, which is possible through use of new age technologies, and already delivered by the authors of this paper as a proof of concept.

- **Empathetic System** with Simplicity to use and Adaptability based on the user: A smart digital census system uses empathy (driven by artificially intelligent libraries) to provide simple systems for data collection. For any self-survey, external tools like Whatsapp or SMS can be used; and for surveys by enumerators, a questionnaire

application that understands the contexts of the enumerator and modifies itself accordingly can be used.

This also ensures that enumerators who may not have any technical expertise can easily use the system.

- **Cloud for Smartness:** A system that is hosted on cloud with high end security features, through a virtual private cloud and strict limitations of usage, the limitations of a digital system can be converted into an advantage.
- **Self-Learning Auditing Technologies:** A smart digital census system is able to constantly improve itself, particularly its auditing methodologies, to increase the accuracy of its outputs. This is possible through use of artificially intelligent algorithms that reduces the need of too much monitoring, as monitoring is done through A.I. that improves itself.
- **A.I. Enumerator Service/ Citizen Service:** When citizens are not able to participate in the census due to certain limitations on their end, they need to approach a central customer service team. When enumerators are facing problems, they need to do the same. For both of their needs, through a clear categorization of potential troubles, a bot can be used for the purpose of such services. The bot is truly smart if it can be self-learning, personalized, and dynamically suited to the needs of the citizens or enumerators of a particular region. This is only possible if an entire ecosystem of smart data collection is used.

This solves a major problem with all types of census, with regards to accessibility by citizens and inefficiency by enumerators. It becomes extremely cost effective, as compared to having an entire team of humans for the purpose, who may not be able to achieve the level of efficiency this system does. We need additional human support here as well for effective outputs.

- **Questionnaire Templates that evolve:** As also seen in the case study covered in this report, there may be special situations in an ongoing survey created due to unknown factors that may require changes in the questionnaire template (for example, COVID-

19 situation that required enumerators to have short surveys due to the guidelines), that needs to be changed and evolved based on inputs from the rest of the survey.

- **A.I. Enumerator Guide:** Imagine a situation where based on predictive analytics (derived from current and past data), an enumerator is recommended the kind of answers expected from a particular region, so that instead of inputting, they need to just modify the responses, because the A.I. puts predicted responses on their behalf.

This reduces the work of the enumerator and makes the entire process easier for someone who may not be used to typing a lot. Of course, in this case, additional monitoring is needed to make sure no false data is getting entered.

- **GPS based tracking of Enumerators:** Enumerators can be guided better to go into specific areas, through intelligent tracking of their movement recorded through the application they carry on their phones. The survey app doubles up as position tracker as well.

With some improvements over a digital census, which is an innovation in itself, a smart digital census system can be created. This smart digital census removes limitations of digital technologies, particularly its acceptance/knowledge, data security, and its accessibility to citizens.

Following were the disadvantages mentioned for digital census that have been explained to be solved by the smart digital census:

- **Comfort with Paper:** This can be solved through involvement of enumerators who are more digital savvy as is the case in the changing world today, and by creating a smart system that is simple enough for enumerators without a lot of education. This has been successfully tried in the case study that is referred to later in this paper.
- **Fear of confidentiality of data:** This is unwanted fear, as use of “cloud” in smart digital census has much more flexibility in terms of the security of data. Also, in order to allay this fear further, multiple combinations of data storage options can be used, depending on the level of confidentiality of data.

- **Increased customer service/technical expertise:** Technical expertise is needed at the level of management of the systems and not at the level of collection of data. But such expertise leads to increased accuracy as the challenges of transferring data from paper to digital do not exist in this case. So, smart digital census replaces hundreds of back-end staff with a few technical staff; with the back-end staff free to focus on other innovative works.

Also, with the conception of a smart digital census, right at the start of the census, technical expertise requirement should not pose a problem. Most importantly, with increased technical expertise, regular customer service requirements will reduce significantly, as much of the customer service is handled by the intelligent systems themselves. This means that the overall costs will also be lesser.

- **Higher Initial Cost:** *A comparison of the expected budgets of a smart digital census vs a traditional census paints a different picture.* In today's world, smart technologies are flexible, requiring almost zero capital expenses (as they can be flexibly handled during operations). There is no expense for developing any system or for having data infrastructure, in a smart digital census, as existing technologies or products can be used on a pay-per-use basis.
- **Lesser Reach to those with low access to technology:** In the smart digital census, this is not a problem at all; as it is designed to suit everyone, including the ones with zero access to technology, since it's the enumerators who use technology and not the survey takers. For the survey takers, reach is even more, because enumerators take less time per household in a smart digital census, thus making sure no one is missed.

So, what remains as the biggest disadvantage of a smart digital census?

The number of resources, both human and systems, used in traditional census is significantly reduced in a smart digital census; and may come across as a replacement of humans by artificial intelligence, though the enumerators' work still remains as earlier. The solution to this disadvantage lies in re-skilling the staff accordingly and preparing them for the future.

8. EVALUATION OF ALL PAPER RESPONSE AND DIGITAL RESPONSE

Table 4: Evaluation of All-Paper Response & Digital Response

Response Technology Areas of Complexity	
Management Complexity	
Data Availability	Complexity associated with the delays in managing operations if data is not available. Quicker data availability enables more timely decision making and better resources allocation while preventing waste (for instance, preventing an enumerator from visiting a household that submitted a questionnaire late).
Data Management	Complexity associated with management of the data for the Decennial and effort required to manage data privacy and security, to prevent data loss, to improve data accuracy, to reduce fraud and to improve data reliability.
Respondent Support	Complexity associated with the need to provide assistance to respondents to support them in successfully complete a questionnaire accurately, including the need to provide access to a proper questionnaire (such as a form in a different language or a duplicate form).
Information Systems Complexity	
Response Capture	Complexity associated with capturing responses from questionnaires for integration into the Census universe, whether provided via self-response, an enumerator, or other means.

Systems Integration	Complexity associated with the number and types of information systems that must be integrated in order to successfully capture, store and manage respondent data.
Organizational Complexity	
Organizational Scale	Complexity associated with managing a number of disparate parties (internal or external) required to support the response technology.
Organizational Experience	Complexity associated with managing new, updated or changed response technology(ies) and supporting systems at an organizational level.

COMPLEXITY OF ALL PAPER RESPONSE

Table 5: Complexity of All-Paper Response

Dimension	Criteria	Analysis	Rating (High/Med/Low)
Management Complexity	Data Availability	An All-Paper response produces the greatest lag time between the questionnaires being completed and the data being integrated into the response universe.	High
	Data Management	An All-Paper response lacks certain data management capabilities that are available in other options. The risk of respondent data loss is highest	Medium

		compared to a digital response due to: ○ Responses lost due to mail or sorting problems ○ Handwriting that cannot be interpreted 100% accurately ○ Stray markings that otherwise result in a misread, preventing complete data capture.	
	Respondent Support	An All-Paper response requires that the Census send forms in an alternate language and/or assist respondents with completing the questionnaire.	Low
Information Systems Complexity	Response Capture	An All-Paper response requires a complex process in order to prepare the paper forms for capture, and then to digitize the data. Forms “read” by software, validated or keyed in by human beings, archived and eventually destroyed.	Medium
	Systems Integration	An All-Paper response requires a relatively moderate level of systems integration, compared to the other two options, due to the need to integrate the DRIS data	Medium

		capture systems with the greater response universe.	
Organizational Complexity	Organizational Scale	An All-Paper response requires cross-functional support from internal departments and external contractors.	Medium
	Organizational Experience	An All-Paper response is highly familiar to the organization and the Bureau already has the experience and processes	Low

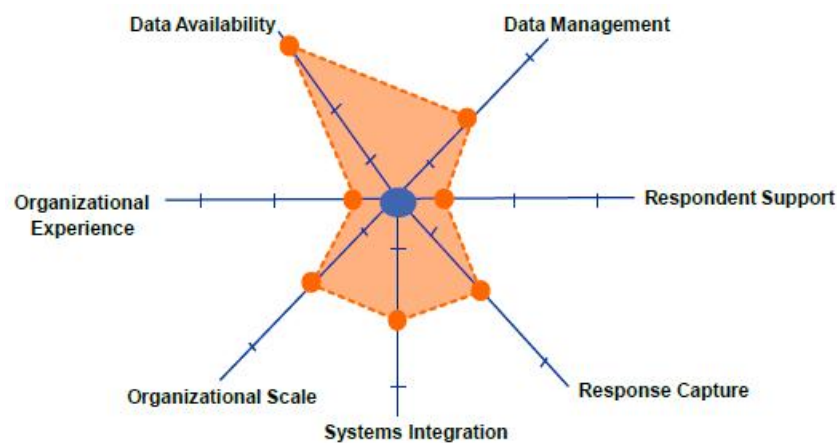


Figure 27: Complexity of all paper response (Source: US Census Bureau)

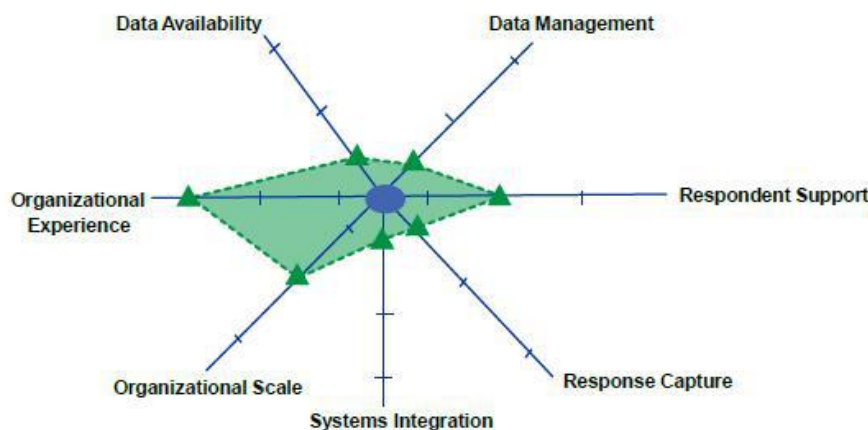


Figure 28: Complexity of digital response (Source: US Census Bureau)

COMPLEXITY OF DIGITAL RESPONSE

Table 6: Complexity of Digital Response

Dimension	Criteria	Analysis	Rating (High/Med/Low)
Management Complexity	Data Availability	A Digital response enables questionnaires completed by all actors to be immediately integrated into the response universe. As a result, respondent data may be made available in real-time for better planning and decision making.	Low
	Data Management	A Digital response provides a high level of support for data management, including safeguards to protect against data loss.	Low

	Respondent Support	A Digital response will require the Bureau to provide a basic level of Technical support for respondents responding via the Internet. “Screen tips” can enable the respondent to receive help on a specific form field right on the questionnaire.	Medium
Information Systems Complexity	Response Capture	A Digital response will enable the Bureau to capture respondent data digitally, eliminating the need to perform a time-consuming and complex digitization. Digital questionnaires enable real-time validation and ensure standardization of data, permitting the instantaneous integration of respondent data into the response universe.	Low
	Systems Integration	A Digital response eases systems integration as a result of the data already being digital and the ability for respondent data to be directly entered into the response universe.	Low
Organizational Complexity	Organizational Scale	A Digital response would likely require a relatively moderate level of organizational coordination in	Medium

		order to implement supporting systems such as firewalls, security, etc.	
	Organizational Experience	A Digital response will require the Bureau to develop new processes and procedures to support the use of this new technology.	High

9. CASE STUDY OF A SMART DIGITAL SURVEY IN THE STATE OF UTTAR PRADESH

Commissioned by the Department of MSME, Government of Uttar Pradesh to UPICO (UP Industrial Consultants Limited), this project undertaken by the team of Sapio Analytics working for UPICO, with support from UFS Limited, was focused on one of the least studied segments that there is, the unorganized sector.

The unorganized sector, due to its sheer nature of being unorganized, has not been studied in such depth; but the challenges for this particular survey were even more:

- **Speed:** Speed of conducting the survey was of paramount importance so that the knowledge gathered by the same can be used immediately for the economic growth of the state that was reeling under the impact of COVID-19.
- **Major changes in the construct of the society and families:** The pandemic brought in a wave of reverse migration to the state of Uttar Pradesh, besides changing a number of aspects about the society, particularly in the households that had members working for the unorganized sector.
- **Limitations in on-ground travel:** With the pandemic looming, door to door traveling was not easy. It is expected that such limitations may become norms at least for the near future, hence innovative methods to handle the same are needed.
- **Fear of bias:** Various reports on the unorganized sectors across the state have resulted out of research that have been functions of assumptions about the unorganized sector. It was important that any such bias did not color the outputs in the wrong manner.
- **Constantly changing situation on ground:** It was realized that the situation on ground was changing regularly, and predictive analytics would be needed to recommend the next steps to the enumerators at regular intervals so that desired data is collected based on the situation. The same was done through use of intelligent auditing, regular secondary research, secondary and primary research based predictions, and regular daily manual feedback to enumerators (who also had the option of taking automated feedback, but it was not applied in this survey).

With all the above-mentioned problems, conducting a survey that can deep dive into the unorganized sector, understand their problems to such levels that hyper-localised policies

around these problems can be recommended, was a clearly a task that would have taken years with involvement of a mammoth task force, if it were not for the Smart Digital Survey that was conducted.

One limitation in this case study is that the survey did not begin with marking of the households against whether they have members or unorganized sector or not, hence leading to outputs that may not be perfect, but are good enough to serve the purpose behind the survey. Such marking before sampling, or full-scale survey of all households (and not just sampling) would have been the best way to conduct the smart digital census for the unorganized sector in the state of Uttar Pradesh under different circumstances.

Let us look at how the above-mentioned challenges were handled for this survey:

- Use of Mobile Applications for data collection in real time
- Use of Algorithms (combining existing Econometric Models with modern Machine Learning Techniques) to derive real time outputs
- Auditing Methods to compare real time outputs with expected outputs, and improve the processes accordingly
- Use of self-surveys (using WhatsApp chat bot, though this was not used extensively) to study the outputs through self-surveys and comparing the same with on-ground surveys to further improve the direction of on-ground surveys, based on required audits
- Enumerator Guidance Systems in the mobile application to help them cover more households in less time
- Use of questionnaire that is simple and less in number to cover larger number of households, while using combination of answers of some of the questions to derive more outputs (this may not be applicable for all such surveys and may have limited scope)

All of the above-mentioned innovative techniques contributed to the creation of **India's first smart digital survey**, and one of the world's first such surveys that combined chat-bot, A.I. engines based on feedback loops, and smart mobile applications; that too for a complicated segment like the unorganized sector, with a unique focus on hyperlocal data.

There was a combination of studies conducted across the districts of Lucknow, Meerut, Gorakhpur, and Varanasi, followed by a *high-end use of various data analytics tools designed*

to analyse the data outputs in the context of each district, created through high end algorithms supported by secondary research and on-groundwork.

Through a deep use of new age technologies and an on-ground workforce, in coming up with various numbers for the unorganized sector spanning multiple demographics and sectoral activities, the report proposed a granular micro-level policy recommendation, that could help enable the unorganised sector to get its fair share of pie in the resource allocation. Resulting it to be streamlined with the main economy of the nation. The core aim of the report was to come up with a well-designed policy recommendation plan for the improvement of the unorganized sector, backed by data driven results, aided by a deep study of policies across the world in similar scenarios.

The report also provided critical data outputs that can be used by various departments for relevant planning and decision making, besides the recommendations generated in the report. The data outputs were accompanied by analysis to help the decision makers, besides being useful for the core policy recommendation.

The team involved in the report included enumerators and their supervisors, responsible for conducting on-ground survey work and collecting data in a pre-planned manner, that was derived based on the work by a team of secondary researchers, who helped create the templates needed by the enumerators for an effective data collection through a mobile app rather than the traditional method of using paper and pen.

This marks the start of a unique concept of digital survey that was done using an Android based mobile application with novel methods of auditing integrated with the same. The data from the survey was made intelligent through a combination of algorithms combined with data analytics tools.

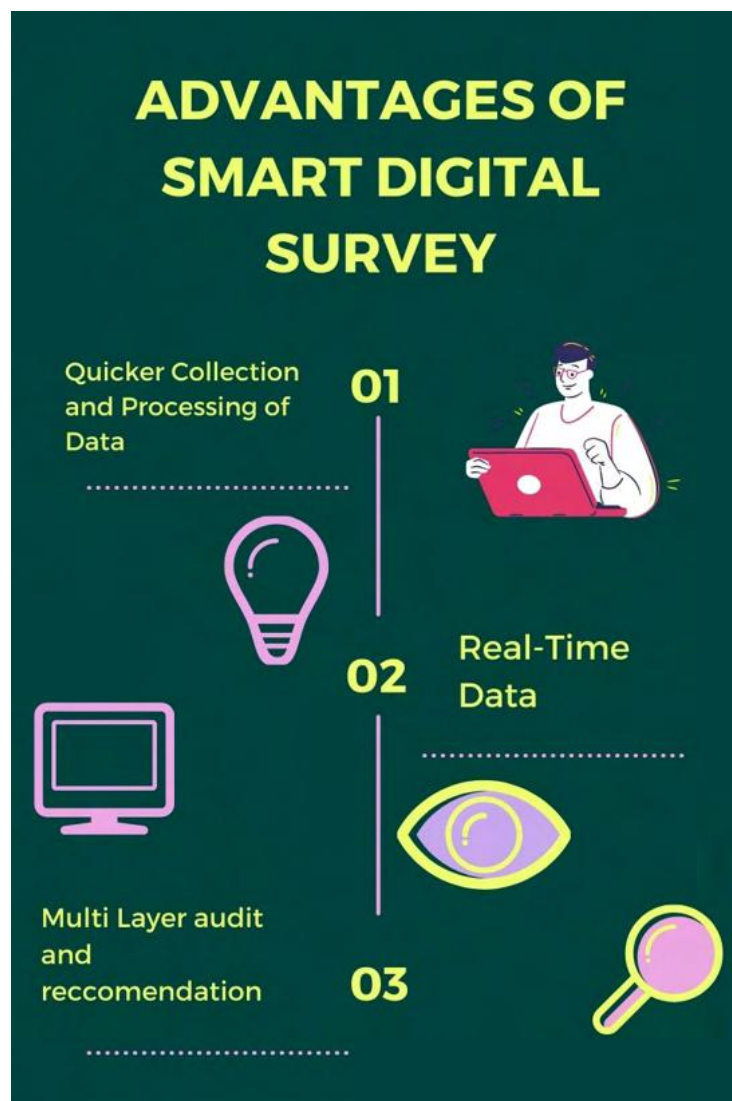


Figure 29: Advantages of Smart Digital Survey

1. Multi-Layer Audit and Recommendations

All the information was received in the digital format as a result of which we were able to create the ecosystem of multi-layer auditing. ***This made us more agile and forward looking.*** Multi-layer audit helped us in analysing the data efficiently and gain insightful knowledge about that region and the possible solutions we could arrive at to combat the situation.

Continuous Auditing of the data enabled us to efficiently analyse data on a frequent basis along with detection of anomalies, outliers, inconsistencies and other factors.

Such multi-layered real time auditing and recommendations of the enumeration process was a function of the following innovative methods:

- Extensive hyperlocal secondary research
- Initial pre-surveys in each location, particularly self-surveys using Survey Chatbots (used on WhatsApp or any other ubiquitous tool)
- Additional self-surveys in each location, to compare outputs and prepare even-better results for newer locations (use of tools like the A.I. Chatbot)
- Manual Audit and daily checks with the enumerators based on automated identification of Triggers (anomalies based on all secondary research and pre-survey reports)
- Enumerator-level performance checks and improvements based on audit in real time

2. Real Time Data

Real time data collection serves the following 3 primary purposes:

- Better auditing and improvement in the processes, as required in the above-mentioned point about multi-layered auditing
- Real time local outputs for actions by the authorities, without waiting for the final report to come out
- Quicker Inputs from Authorities on any changes needed in the fieldwork (questionnaire, etc).

It allowed to make quick and responsive changes to survey plans and enhancing the depth and breadth of fieldwork without expending extra time and effort. Availability of real time data *enabled us to enhance our degree of reliability on the quality of data collected hence mitigating the risk of wrong collection of data and ensuring that the sample selected reflected the true population.*

Real-time analytics helped us pinpoint issues the moment they arise, and in some cases, catch them before they occur.

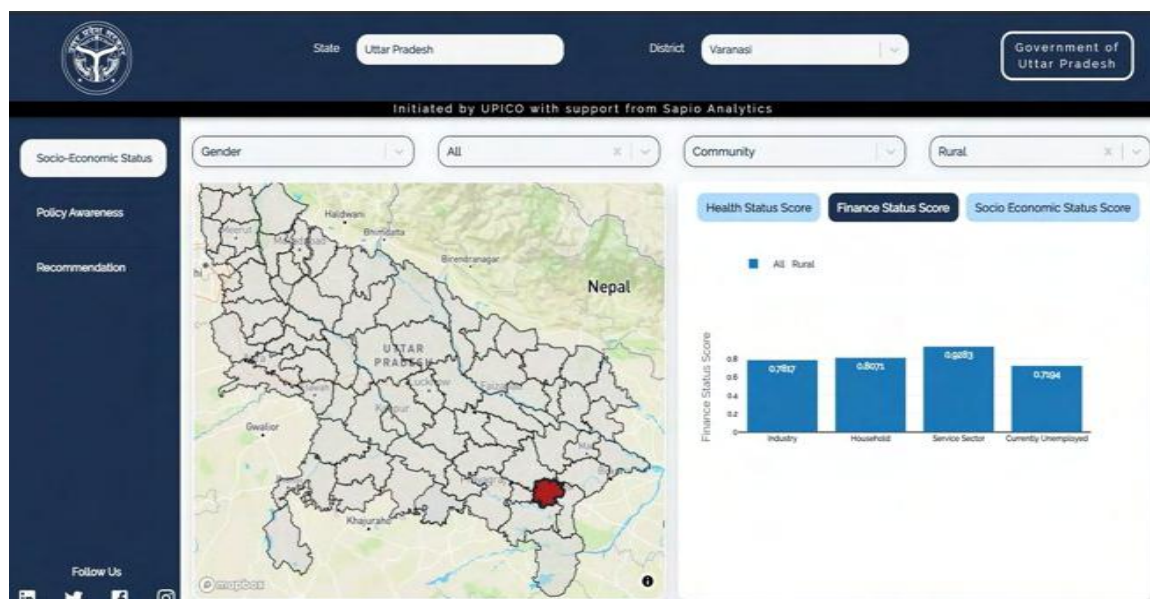


Figure 30: Sample screenshot of real time dashboard

This real time nature of the survey was achieved through the following innovative methods:

- Data stored on cloud: Since mobile applications were used, no data was saved on the local handset; all data went to cloud in real time, allowing the processes to move faster and skipping any need on the side of enumerator to take care of the data collected by them; as it was all system driven.
- Use of algorithms to do calculations on data in cloud to give real time results in form of outputs
- Easy and flexible systems/applications: Mobile applications come with administrator dashboards to make real time changes that reflect immediately on the applications.

3. Quicker Collection and Processing of Data

Our smart digital survey enables the enumerator to have a pre-determined questionnaire, recommended based on their location and other demographic parameters as instructed to them, with certain known fields already filled. This saves time and helps in quicker collection of data with a smaller number of enumerators used.

For example, if an enumerator is supposed to visit locality X in a particular district, based on the pre-defined intelligence in the system, they will get a few templates mapped to the locality

X with pre-filled answers that can be edited easily. So, they can choose the templates and make edits rather than filling everything from scratch.

This is one of the ways in which quicker collection happens. Following are the various methods used in a smart digital survey system that can be used:

- Enumerator Guidance using Templates (as mentioned above)
- Regular Enumerator Recommendations (automated)
- User friendly data collection system
- Offline data collection in absence of internet: We ensure that internet connectivity is never a limitation, as data can be collected offline and sink when the device used for collection comes offline.

Through use of the above-mentioned methods, data collection is easy and fast, becoming one of the trademarks of the smart digital census.

The figure shows four sequential screens of a digital survey application. Each screen has a back arrow and a user icon at the top.

- Screen III: OCCUPATION**

Where do you work?

 - a. Service Sector ☐
 - b. Industry ☐
 - c. Households ☐
 - d. Currently Unemployed ☐
 - e. Doesn't want to ans ☐

Next
- Screen TYPE OF EMPLOYMENT**

1. What type of respondent?

 - a. Self Employed ☐
 - b. Owns Business ☐
 - c. Contractual work for other business ☐
 - d. Doesn't want to answer ☐

2. Is your workplace a proper set-up?

 - a. Yes ☐
 - b. No ☐
 - c. Doesn't want to ans ☐

3. Do you have multiple sources of income (apart from the core work activity being done by you)?

 - a. Yes ☐
 - b. No ☐
 - c. Doesn't want to ans ☐
- Screen II: PERSONAL IDENTIFICATION**
 - 1. Aadhar Number
 - 2. Mobile Number
 - 3. Registration Number
 - 4. UAN
 - 5. ESIC
 - 6. Name
- Screen IV: HOUSEHOLD P&L**
 - 1. Total Monthly Income in the Household (including income from cattle or land)
 - 2. Total Monthly Expenses of the Household
 - 3. Total Cumulative Savings :
 - Cash in Hand/Bank
 - Jewelry
 - Others
 - 4. Fixed Asset
 - ☐ Land
 - ☐ House
 - ☐ Cattle

Figure 31: Sample screenshot of the application

Below is the table showing the time and the manpower used for conducting the survey in all the regions, in this particular case.

Table 7: On-Ground Survey for Uttar Pradesh

Location	No. of Enumerators	Time
Lucknow	52	25 days
Meerut	25	20 days
Gorakhpur	25	12 days
Varanasi	51	66 days*

*Additional time was taken here due to certain unforeseen technical circumstances and should be considered an anomaly. If this would have been the all-paper response method, it would have far more enumerators and days to complete the survey.

4. Effective Visualization for Quick Actions

A smart digital census system also believes in the power of smart visualisation of various outputs, that can be used for quick grasp of the situation, based on which immediate actions can be taken.

“If you have to show the seriousness of a situation in just a few seconds, use of a visual heat map focused on colours without numbers may do the trick.”

~ Shri. Jayant Kumar Banthia, Census Commissioner of India from 1999 to 2004 (oversaw the 2001 census), former Chief Secretary, Government of Maharashtra.

In our interaction with Mr Banthia, we understood the importance of visualisation, and realised that smart digital census can't be created without a focus on visualisation, that is able to speak to the decision makers in just a few seconds if required.

Use of graphs along with heat maps (using proprietary GIS technologies) for this case study helped in all aspects of this study. In future studies, additional visualisation can be done in the form of 3D GIS.

10. BENEFITS OF SMART DIGITAL CENSUS

This paper focuses on a number of benefits that smart digital census provides over any other traditional form of census, or even the conventional digital census.

Here is a summarised understanding of the said benefits:

- **Financial Benefits**

When seen from the context of the case study (survey done in the state of Uttar Pradesh for the unorganised sector), this benefit comes out very clearly.

A much lesser number of enumerators can cover a large population, as compared to a traditional census, and in a much lesser time. Both the number of people and the time taken are effectively the major costs involved in the process of data collection.

In addition to the same, cost is further reduced through absence of later post-processing work due to the real time nature of the same. Of course, the technology costs are additional, but their numbers are much lesser than the benefits in cost due to the technology.

- **Social Benefits**

Following clear benefits are there when a smart digital census is conducted:

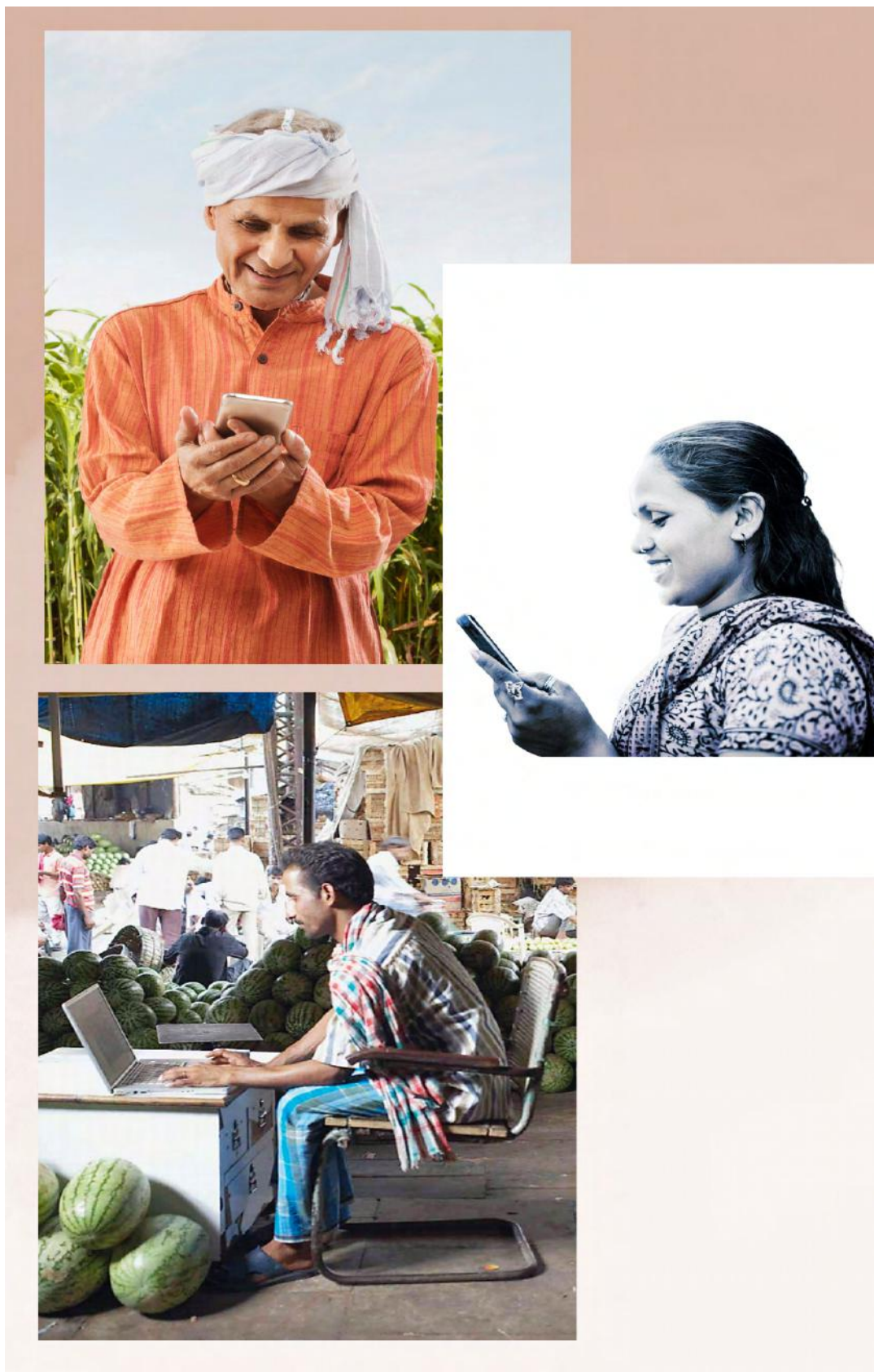
- More personalized interactions with the citizens: Technology enables the interactions to be more personalized, especially the technologies around the enumerator guidance. This makes sure that people are involved and helps in reducing inaccurate information.
- Feeling of contribution to change: Since the process is real time and provide outputs very quickly; if the authorities choose to publish parts of the same, the citizens feel more involved in contributing to the change, and this participation leads to ensuing social benefits.
- Increased reach to all households: The traditional census also excels in a very widespread reach. But with the power of digital technologies making the outreach easier, the households hidden from traditional census are also expected to be covered through a smart digital census, that uses hyperlocal reach to make sure no segment is missed as per the requirements of the census.

- **Administration Benefits**

The benefits for the administration, in terms of increased efficiency and accuracy have already been mentioned multiple times in this paper. The administration sees better results in less time, with intelligent recommendations, at a reduced cost, and particularly useful in situations like the current pandemic where dependency on digital technologies is higher.

One of the major outputs of adoption of smart digital census is Increase in Frequency of census across demographics and segments, because of its ease in conducting the same, leading to a much better understanding of the problems of India's citizens.

11. WHY IS INDIA READY FOR CHANGE?



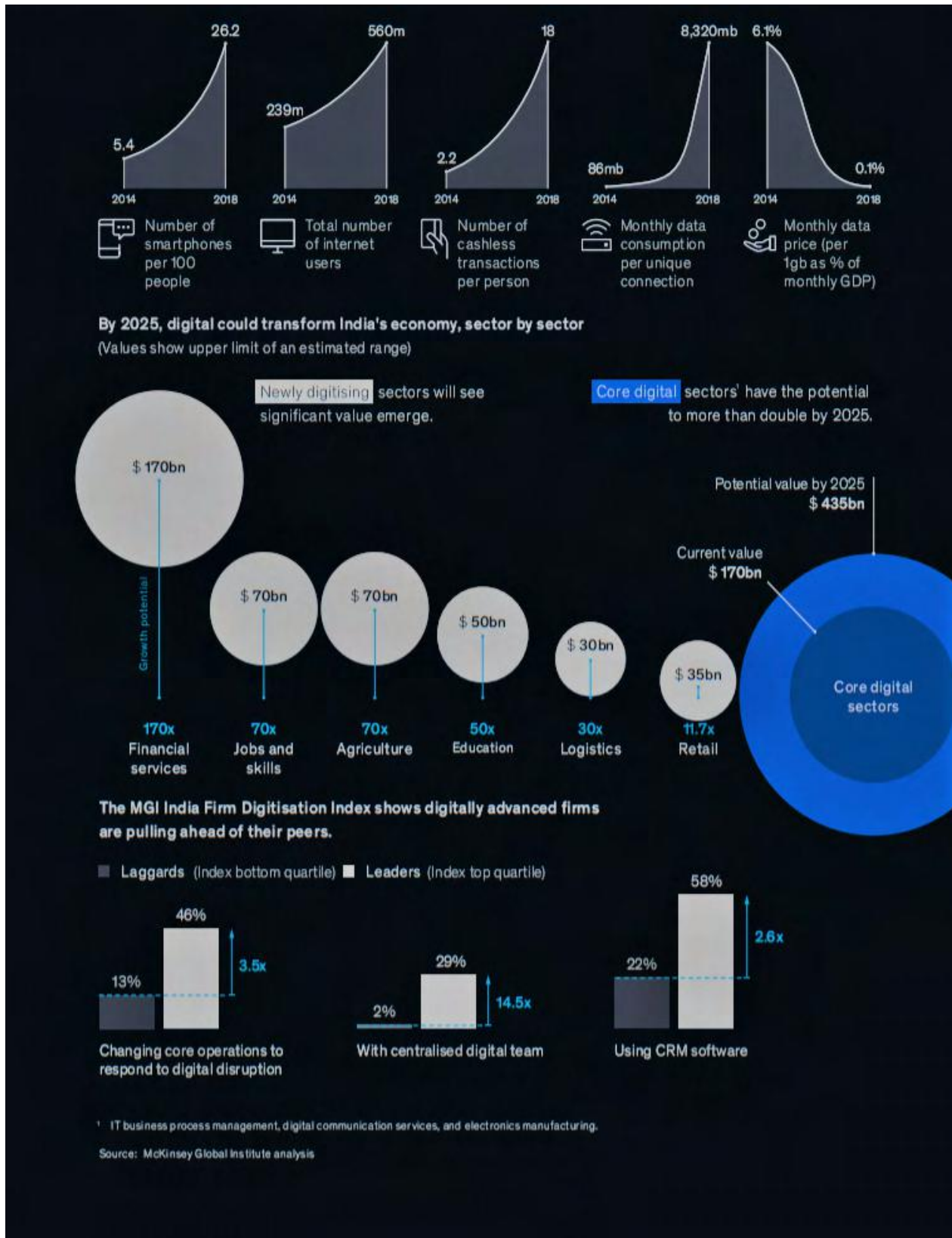


Figure 35: Unlocking the potential of technology in India (Source: McKinsey Global Insitute Analysis)

India is among the top two countries globally on many key dimensions of digital adoption.

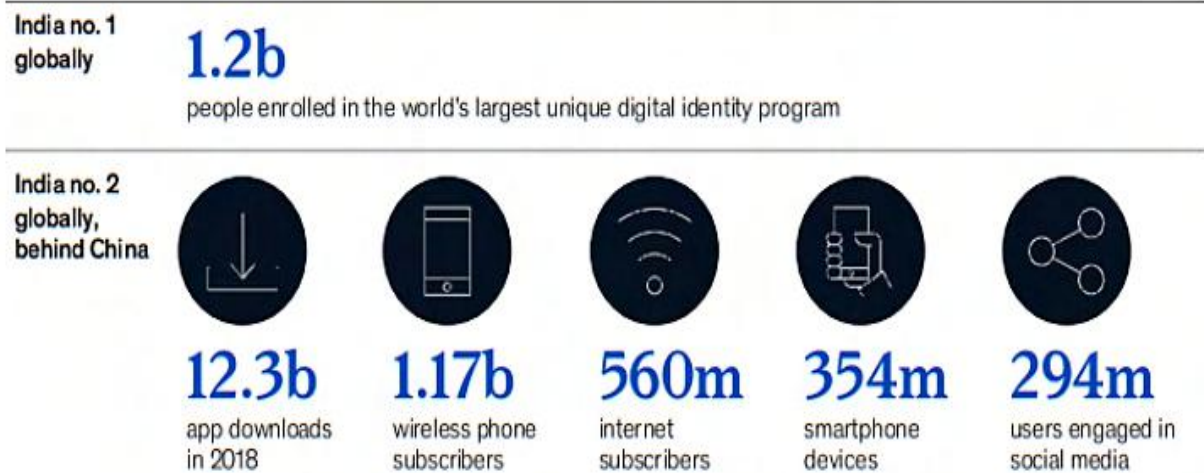


Figure36: India's digital adoption at a glance (Source: Priori Data, Strategy Analytics; Telecom Regulatory Authority of India, Unique Identification Authority of India)

With more than half a billion internet subscribers, India is one of the largest and fastest growing markets for digital consumers, and the rapid growth has been propelled by public and private sector alike. India's lower-income states are bridging the digital divide, and the country has the potential to be a truly connected nation by 2025. Much more growth is possible. As India's digital transformation unfolds, it could create significant economic value for consumers, businesses, microenterprises, farmers, government, workers, and other stakeholders.

India's internet user base has grown rapidly in recent years, propelled by the decreasing cost and increasing availability of smartphones and high-speed connectivity, and is now one of the largest in the world. The country had 560 million subscribers in 2018, second in the world only to China. *Digital services are growing in parallel. Indians now download more apps—12.3 billion in 2018—than residents of any other country except China.* The average Indian social media user spends 17 hours on the platforms each week, more than social media users in China and the United States. *The share of Indian adults with at least one digital financial account has more than doubled since 2011, to 80 percent, thanks in large part to the more than 332 million people who opened mobile phone-based accounts under the government's Jan-Dhan Yojana mass financial-inclusion programme.*

The public sector has been one strong catalyst for India's rapid digitisation. The government's effort to ramp up Aadhaar, the national biometric digital identity programme, has played a major role. The Goods and Services Tax Network, established in 2013, brings all transactions involving about 10.3 million indirect taxpaying businesses onto one digital platform, creating a powerful incentive for businesses to digitise their operations.

At the same time, *private-sector innovation has helped bring internet-enabled services to millions of consumers and made online usage more accessible.* For example, Reliance Jio's strategy of bundling virtually free smartphones with subscriptions to its mobile service has spurred innovation and competitive pricing across the sector. Overall, *data costs have dropped by more than 95 percent since 2013: the cost of one gigabyte fell from 9.8 percent of per capita monthly GDP in 2013 (roughly \$12.45) to 0.37 percent in 2017 (the equivalent of a few cents).* Average fixed-line download speed quadrupled between 2014 and 2017. *As a result, monthly mobile data consumption per user is growing at 152 percent annually—more than twice the rates in the United States and China.*

12. CONCLUSION

Population and housing census provide a wide range of statistical information on the population – persons, households and housing units. They allow detailed disaggregation of data by small geographic areas and small population groups. However, it is widely recognised that conducting a population and housing census is one of the most expensive and complex data collection operations, comprising of a series of many interrelated activities. The logistical effort required, and the difficulties encountered in contacting and collecting information on the whole population within a limited period, adds to the complexity.

Governments are increasingly facing several challenges in conducting population and housing censuses, ranging from falling participation to increasing data collection costs. Response rates have declined as citizens become increasingly concerned about information security and privacy and the confidentiality of information given to the government. Households are becoming more diverse, dynamic and mobile, making it a challenge to reach people and assign them to a single unique location. These societal, demographic and budgetary trends are making it harder and more expensive to locate individuals and solicit their participation through traditional data collection methods. Moreover, each of these trends threatens the ability of government to deliver high-quality, timely and cost-effective census results.

These challenges are compelling Governments to investigate alternative ways of implementing the census, and ultimately, to modernise and transform the traditional method of conducting a census. Today's information and communication technologies offer unprecedented opportunities to innovate and transform their census operations, dramatically improving performance and census results while reducing costs. Moreover, they can enhance the quality and accessibility of census results. With proper planning, governance and vision, modern technology can help to improve the efficiency (as measured in terms of cost and time savings, productivity gains, increased accuracy) and effectiveness (by, for example, providing better products and service, improved analysis) of census processes.

Recent advances in information and communication technology have changed how routine statistical business processes such as Enumeration Area (EA) design, data capture and validation, data processing, and data dissemination, are being carried-out. The use of such technology has become an integral and vital part of census processes critical for improving the

cost, quality (coverage, accuracy, timeliness) and efficiency of the census. The rapid expansion in mobile connectivity, and rapid progress in technological innovation more broadly (such as cloud computing, smart mobile devices, GPS, web GIS), provide new opportunities for increasing the quality and speed with which census data can be collected and the statistics produced. Advancements in technology has brought digital media in forefront of everyday life. Citizens of India are becoming increasingly comfortable with digital modes and it is expected to grow even further. We are told time and again that data is a prized commodity, but it is very easy to be big data rich, yet still insight poor, if information isn't harvested and analysed efficiently.

Hardware and software as well as mobile data networks are becoming cheaper and better, and infrastructure and capacity that enables effective deployments is increasingly available in all parts of the world. This means that the immense progress in connectivity (in terms of both speed and magnitude) and technology adoption has a large potential for modernizing and complementing traditional statistical collection. New approaches are allowing the collection of data with handheld electronic devices, the Internet and the telephone in a manner that is better, faster and potentially cheaper. It's imperative to say that DIGITAL CENSUS is the way forward for the inclusive development of society as whole. It will benefit both public and private sector and will display the perfect model of public-private partnership.

----End of Report ---

13. REFERENCES

- *India Today* (2021), Census to go digital for the first time, February 1, retrieved from:
<https://www.indiatoday.in/business/budget-2021/story/budget-2021-census-to-go-digital-for-the-first-time-with-rs-3768-crore-allocation-nirmala-sitharaman-1764702-2021-02-01>
- *The Conversation* (2017), Welcome to E-Estonia, the tiny nation that's leading Europe in digital innovation, April 4, retrieved from:
<https://theconversation.com/welcome-to-e-estonia-the-tiny-nation-thats-leading-europe-in-digital-innovation-74446>
- Raghavan, P. (2010), Brief History of Census in India, *Financial Express*, April 2, retrieved from:
<https://www.financialexpress.com/archive/column-brief-history-of-census-in-india/598794/>
- Stokes, Rose. (2020), The implications of the first digital census, *Raconteur*, December 2, retrieved from:
<https://www.raconteur.net/public-sector/implications-digital-census/>
- Ministry of Home Affairs, *Census Operations*, retrieved from:
https://censusindia.gov.in/Data_Products/Library/Indian_perceptive_link/Census_Operation_link/censusoperation.htm
- Kumar, Vikas. (2020), Census 2021 will be delayed. It gives Modi govt time to bring long-pending reforms, *The Print*, August 13, retrieved from:
<https://theprint.in/opinion/census-2021-will-be-delayed-it-gives-modi-govt-time-to-bring-long-pending-reforms/479896/>
- Zachary, Schwartz. (2020), U.S. Census Bureau: Digital transformation of the 2020 Census, September 23, retrieved from:
<https://www.pega.com/customers/us-census-bureau-digital-platform>

- KPMG India. (2019), Digital India – Power to Transform, June 28, retrieved from: <https://home.kpmg/in/en/home/insights/2019/06/digital-transformation-india.html>
- Money control (2019), Amit Shah proposes multipurpose ID card: What is it and how feasible is it for India?, September 25, retrieved from: <https://www.moneycontrol.com/news/india/amit-shah-proposes-multipurpose-id-card-what-is-it-and-how-feasible-is-it-for-india-4475271.html>
- McKinsey Global Institute (2019), Digital India: Technology to transform a connected nation, March, retrieved from: <https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Digital%20India%20Technology%20to%20transform%20a%20connected%20nation/MGI-Digital-India-Exec-summary-April-2019.pdf>
- Zhou Youyou (2019), Ghana is betting on a digital census in 2020 to transform its informal economy, Quartz Africa, retrieved from: <https://qz.com/africa/1630207/how-ghana-conducts-its-2020-digital-census/>
- Thieme Michael (2019), Considering a Digital 2020 Census, U.S. Census Bureau, retrieved from: https://nces.ed.gov/FCSM/pdf/Thieme_2012FCSM_IV-B.pdf
- Chartered institute of internal auditors (2019), Keep it real: what is real time auditing?, September, retrieved by: <https://www.iaa.org.uk/audit-risk-magazine/features/keep-it-real-what-is-real-time-auditing/>
- Howells Jan (2018), How much data is too much?, Orange Business Services, retrieved from: <https://www.orange-business.com/en/blogs/how-much-data-too-much>
- Whitler Kimberly (2018), Why Too Much Data Is A Problem And How To Prevent It, Forbes, retrieved from: <https://www.forbes.com/sites/kimberlywhitler/2018/03/17/why-too-much-data-is-a-problem-and-how-to-prevent-it/?sh=6972273b755f>

- United Nations Statistics Division (2020), Report on the results of the UNSD survey on 2020 round population and housing censuses, March 6, retrieved from:
<https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3j-Survey-E.pdf>
- Kemp Simon (2021), Digital 2021: India, Datareportal, retrieved from:
<https://datareportal.com/reports/digital-2021-india>
- Demirci Meryem (2020), United Nations Statistics Division Programme in Support of the 2020 Round of Population and Housing Censuses and Civil Registration and Vital Statistics, United Nations Statistics Division, retrieved from:
https://www.unescap.org/sites/default/files/Session2_Programme_support_2020_Round_Population_Housing_Censuses_UNSD-Workshop_SDG_PHC_CRVS.pdf
- Guidelines on the use of electronic data collection technologies in population and housing censuses (2019), Department of Economic and Social Affairs, retrieved from:
<https://unstats.un.org/unsd/demographic/standmeth/handbooks/data-collection-census-201901.pdf>