

SUE Survey Planning Team Guide

Surveys for Urban Equity (SUE) Project

Version 1.7

Authors:

Dana R. Thomson

Radheshyam Bhattarai

Rajeev Dhungel

Subash Gajurel

Sushil Singh

Shraddha Manandhar

Sudeepa Khanal

Acknowledgements: Special thanks to the HERD International ICT Department for help preparing the IT Team instructional materials, and Sushil Baral, Mashreky Saidur, Silvia Junnatul Ferdoush, Tarana Ferdous, Duong Minh Duc, Hoang Van Minh, Nguyen Bao Ngoc, Ak Narayan Poudel, Helen Eley and the rest of the SUE Project Team for providing critical input to this work. Also thanks to Bijay Dhakal and Raziv Pathak for preparing instructions to update OpenStreetMap with JOSM.

Recommended citation: Thomson DR, Bhattarai R, Dhungel R, Gajurel S, Singh S, Manandhar S, Khanal S. 2018. Surveys for Urban Equity (SUE) Project: Planning Team Guide. Leeds: Leeds University. 165 p.

Available at: <https://medhealth.leeds.ac.uk/info/691/research/2388/sue>.

Contents

1. Introduction.....	4
About Surveys for Urban Equity (SUE)	4
Motivation.....	5
2. Definitions	9
Typical household surveys	9
Typical household survey: workflow	12
Typical household surveys: potential problems	13
SUE surveys.....	16
SUE Survey: Workflow	19
SUE surveys: addressing potential problems	20
SUE surveys: evaluation.....	21
3. Workflow	24
Overview.....	24
Survey design.....	27
Mapping and listing	27
Interviewing	28
4. Survey Design and Selection	31
Sample size, design, budget	31
GridSample.....	34
Sample frame scale.....	34
Defining PSUs	35
5. Infrastructure & Tools	39
Server.....	39
Set-up of tablets, computers, and forms.....	39
GIS Workstation	41
Tablets	42
6. Logistics: Staff	46
Hiring.....	46
Training.....	47
7. Logistics: Pre-field.....	49
Set-up tools and infrastructure	49
Determine final PSUs	49
Team formation.....	55
Pre-field segmentation	56

8. Logistics: Mapping-Listing Fieldwork.....	58
Quality Control	58
Office meetings.....	65
Field visits.....	66
9. Logistics: Post-field	68
Select household sample (two-stage only)	68
Post-field segmentation (one-stage only).....	69
Download permanent record of OpenStreetMap data.....	70
Map shack/tent-residential buildings offline.....	70
Prepare Map Atlas for interviews.....	70
Prepare tablets for interviews.....	71
Using PSU maps & dwelling/household listing for interviews.....	72
Sample Weights.....	73
References	78
Appendix A: Exercises	79
Exercise: Reproduce Nepal SUE survey sample size calculation.....	80
Exercise: Become familiar with the GridSample website.....	82
Exercise: Sample frame scale	84
Appendix B: Forms	88
Training agenda.....	89
Mapper-Lister Job Announcement	91
GIS Job Announcement.....	93
Planning team OFFICE tracking form	95
Planning team FIELD tracking form.....	97
Office mapping tracking form	99
Tracking Sheet: GIS Team	100
Tracking Sheet: Mapping-Listing Team.....	102
Detailed mapping-listing schedule.....	103
Example mapping-listing field exam map and answer key	104
Mapping-listing field exam score sheet	106
Daily training feedback form	108
Appendix C: How-Tos for IT team	109
How To: Setup GoeODK server	110
How To: Setup tablets	121
How To: Setup GeoODK form in Excel.....	149

How To: Pull listing data into questionnaire form	152
How To: Complete one-stage interview questionnaire	156
How To: Complete two-stage interview questionnaire.....	161
Appendix D: Exercise Answers	164

1. Introduction

Objectives:

By the end of this chapter, you will understand:

- Purpose of the Surveys for Urban Equity (SUE) project
- Reasons to test survey new methods and tools

Outline:

- About Surveys for Urban Equity (SUE)
- Motivation
- New SUE methods
- New SUE tools

About Surveys for Urban Equity (SUE)

The **Surveys for Urban Equity (SUE)** is a research project coordinated by University of Leeds (UK) in collaboration with HERD International (Nepal), ARK Foundation (Bangladesh), CIPRB (Bangladesh), Hanoi School of Public Health (Vietnam), and Southampton University's WorldPop-Flowminder-GridSample team (UK). The SUE Project uses a mix of qualitative and quantitative methods to improve the representation of the urban poor in household surveys, and to improve the use of survey data by local authorities.

The SUE Project involves three face-to-face household surveys in Kathmandu (2017), Dhaka (2018), and Hanoi (2018). The Kathmandu survey is the largest with 1200 households in 60 sampling areas, and it is the site where new survey methods and tools were tested and evaluated. The Dhaka and Hanoi surveys are each smaller with 400 households in 20 sampling areas to collect comparative data and evaluate how well the survey methods and tools developed in Kathmandu transfer to other urban contexts.

See Figure 1 for an overview of the SUE project main activities. The activity of piloting gridded population sampling and OpenStreetMap-based enumeration aims to improve representivity of the urban poorest in household survey data. By using the robust methods and tools outlined in this manual to map buildings and list households/dwellings, we support the next project activity to pilot household survey questions about injuries and mental health which are highly relevant to the health of the urban poorest, so that new questions might be included in future national surveys. The ultimate aim of the SUE project is to reduce urban health inequities.

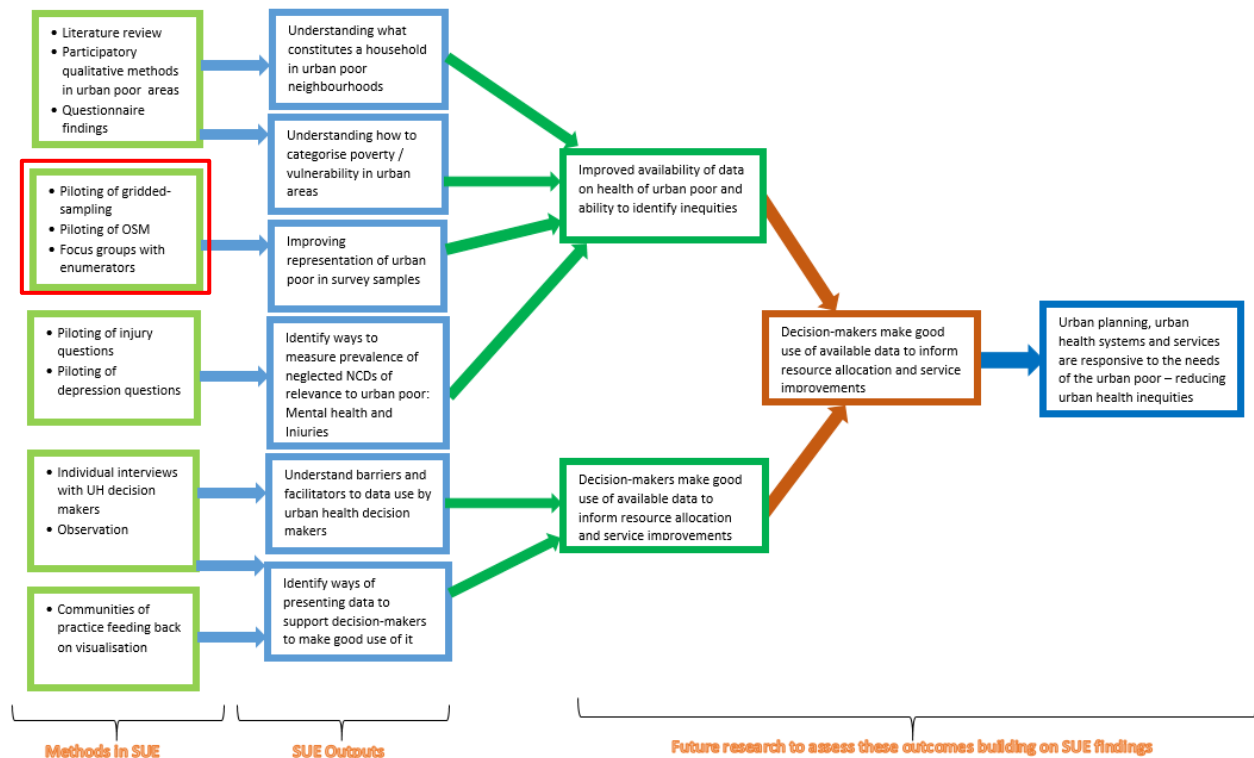


Figure 1. Surveys for Urban Equity (SUE) Project Conceptual Framework

Motivation

Large-scale household surveys are an essential source of data about the health and socioeconomic status of populations around the world. Standardized surveys, such as the Demographic and Health Surveys and Multiple Indicator Cluster Surveys, interview approximately 10,000 households every five years in low-income and middle-income countries around the world. It is because of these surveys that we know national rates of childhood diarrhoea, vaccination coverage rates, childhood mortality, and many other health indicators.

Surveys are essential because they include people who do not visit health centres due to social, economic, or environmental barriers, and thus do not contribute data to the health system. In places where people have good access to the health system, surveys are still helpful because national health systems have difficulty aggregating data from a mix of public and private institutions. Because millions of pounds of government and international money is targeted to vulnerable regions and groups based on national survey statistics, it is imperative that surveys are representative of populations, including the poorest.

The urban poorest are likely under-represented in household surveys because **household survey methods were established four decades ago and have not changed much while populations have undergone major social-economic transitions**. Typical survey methods were designed for predominately rural populations before the internet, satellite imagery, global position systems (GPS), or “big data” were widely available. When survey methods were established, official census data were essentially the only main source of population information to select representative household samples.

In recent years, most survey implementers have switched from paper-questionnaires to tablet-questionnaires. However, enumeration and listing is still largely paper-based. Generally, buildings in sampled communities are identified with hand-drawn paper maps, and a list of households is collected on paper.

Societies have changed dramatically over the last 40 years. In 2007, the earth's population became majority urban. Technology changes have facilitated rapid globalization of ideas and products which have shifted economies, diets, exercise and smoking habits, especially among urban populations. Mobility has also accelerated. More and improved roads, and more and cheaper flights makes seasonal work and living away from family more common today than 40 years ago. Nowhere are these realities more familiar than in south Asia.

The SUE Project is evaluating the following methods and tools to improve the accuracy and feasibility of household surveys in complex urban environments.

New SUE methods:

- **One-stage sampling:** Typical household surveys use a two-stage sample design in which enumeration areas (EAs) from the last census are sampled first, then households are listed and sampled. Two-stage sampling is a sensible design when the population is clustered in homogenous villages or neighbourhoods, and when EAs (approximately 400-600 people each) are sampled in a first stage.

However, two-stage sampling requires identification of the same households twice, several months (or even years) apart, making mobile households vulnerable to exclusion and non-response. Furthermore, conducting a full listing of households in EAs is time- and resource-intensive. One-stage sampling is an attractive alternative in complex urban environments because household listing and interviewing can occur on the same day, dwellings rather than households are listed (removing the need to identify head of household name), and mapping-listing covers a much smaller area which greatly reduces fieldwork.

Household surveys are no longer tethered to census EAs for sampling and modern cities are highly heterogeneous. In Kathmandu, for example, a single building commonly houses a wealthy household on the top level, middle-class households in the middle levels, and poor households in crowded conditions on the bottom levels. Informal housing structures are constructed throughout the city, and not necessarily concentrated in large slums. In this complex urban environment, the design effect introduced by similar households being located next to one another is not expected to be much larger in one-stage surveys than two-stage surveys, and can be overcome by increasing the number of sampling areas.

The SUE Project will compare the feasibility and types of households sampled in one-stage versus two-stage sample designs. The two-stage sampling areas will have approximately 200 households each, and the one-stage sampling areas will have approximately 20 households each.

- **Loosening the standard definition of household:** Large-scale household surveys define a “household” differently, and these definitions were initially shaped by rural contexts in past decades. While definitions have evolved to accommodate more complex living arrangements in cities, standard definitions of a “household” still exclude certain types of people, for example, a cleaner who sleeps at an office building in the city most nights and only occasionally visits her family home outside of the city. This loosening of definition is a step toward the measurement of “individuals” rather than “households” in cities, which has been recommended for measurement of urban populations (Campbell 2017). Note, this approach still allows households to be identified and analysed in the data.
- **Vertical mapping:** Typical household surveys map buildings in space, but building levels are not recorded. In dense urban settings, dozens, even hundreds, of households can occupy the same building. Thus the SUE Project will incorporate vertical listing, recording the level on which each dwelling is located. Vertical listing is particularly important for segmenting one-stage sampling areas in high density areas.

New SUE tools:

- *GridSample* – GridSample is used to select a representative sample of primary sampling units based on WorldPop gridded population datasets. WorldPop estimates total population in 100m X 100m grid cells, and reflects 2015 urban extent boundaries; 2017 buildings, residential areas, places of interest, and roads from OpenStreetMap; and 2010 landcover types detected by satellite imagery.

- *OpenStreetMap* – OpenStreetMap is leveraged as a source of existing geographically accurate data. We contribute to this public data source during the enumeration process, benefitting a multitude of humanitarian and health applications that depend on OpenStreetMap data.
- *GeoODK* – GeoODK is used for all field data collection forms including the dwelling listing form (one-stage sampling areas), household listing form (two-stage sampling areas), and household and individual questionnaires.

Review

Answer the questions below. To see if your answers are correct, go to Appendix D.

1. What do we call a new method for listing a dwelling by its building location and level?
2. What tool will be used to collect dwelling and household listing information?
3. What are reasons to consider one-stage sampling (with gridded population data) over two-stage sampling (with census EA data)?
 - a. Possibly more accurate listing and sample of mobile and vulnerable populations.
 - b. Reduced time and cost of mapping-listing activities.
 - c. Both a and b.

2. Definitions

Objectives:

By the end of this chapter, you will be familiar with:

- Terms used to design and implement surveys
- Typical survey methods and tools
- SUE survey methods and tools, and why they are used

Outline:

- Typical household surveys: terms
- Typical household surveys: workflow
- Typical household surveys: problems
- SUE surveys: terms
- SUE surveys: workflow
- SUE surveys: addressing problems

Typical household surveys

To understand the typical household survey workflow, let us first review a number of terms.

Representative household survey: A face-to-face survey designed such that responses are representative of the population, and where households are the final sampling unit.

Sample: In the context of household surveys, sampling is always “randomized”; “Purposeful” sampling is never used in household surveys. If certain sub-groups are important to measure (e.g. slum areas), then the survey designer can stratify or oversample by sub-group, but this requires that sub-group membership is known in the sample frame (e.g. sampling units are classified as slum or non-slum).

Administrative areas: Official government boundaries of areas. For example wards, municipalities, districts... up to national border. Census enumeration areas (EAs) of 400-600 people are commonly sampled in household surveys.

Sample frame: An exhaustive list of small areal units, usually administrative areas or enumeration areas from the last census, covering the population of interest. The total population per unit should be noted if units differ substantially in population size.

Probability proportionate to (estimated) size: If all of the units in the sample frame have roughly the same population, then a simple random sample of units will result in a sample representative of the population. However, if units in the sample frame have different population sizes (which is almost always the case), then units must be sampled with probability proportionate to size. If the population sizes are known (e.g. from a recent census) then the sample selection is PPS. If the population sizes are estimated (e.g. population totals are projected from an old census) then the selection is PPES.

Cluster: A group of households that comprise a small area, such as an administrative unit. Often refers to a PSU (see below). Cluster and PSU are often interchangeable terms.

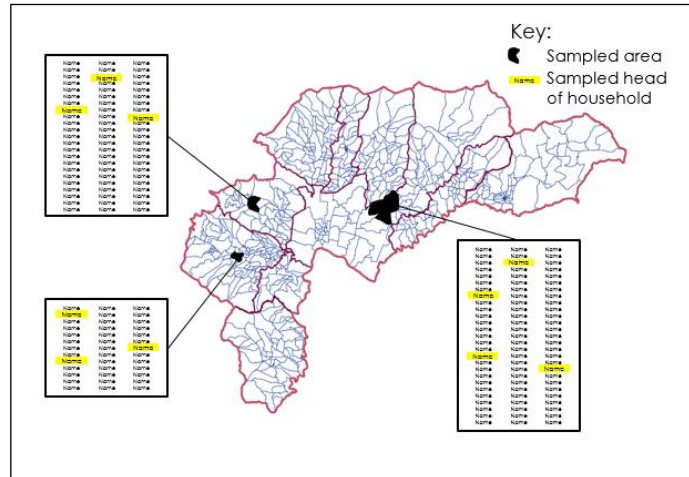


Figure 2. Depicts a typical two-stage cluster sample design - list administrative areas, sample administrative areas, list households, sample households

Multi-stage cluster sampling:

“Multi-stage” sampling means that sampling occurs more than once during the sample process. “Cluster” sampling means the sampling units are embedded such that the first unit (administrative areas) is a group of the next sample unit (households).

Primary sampling unit (PSU): The result of the first stage of sampling (usually a census enumeration area). By the same logic, a secondary sampling unit (SSU) is the result of the second stage of sampling (usually a household in typical household surveys).

One-stage (cluster) sample: This means that we sample small areas, list all households in the sampled areas, and then interview all of the households in the sampled areas. Also called a “micro-census”.

Two-stage (cluster) sample: This means that we sample small areas, list all households in each sampled area, sample households from the list, and then interview the sampled households. This is the typical design of large-scale representative household surveys.

Stratification: Exhaustive sub-groups of the population, for example all provinces, or both/all genders. Stratification means that independent samples are selected in each sub-group (called a stratum). Often equal numbers of households are sampled in each stratum, though sometimes proportional or statistically optimal numbers of households are sampled.

Oversampling: Oversampling of domains is sometimes performed alongside stratification. Oversampling ensures that the minimum sample size (per stratum) is met for each domain in the overall combined-strata population. Extra –PSUs/clusters are selected in the under-represented domain.

Segmentation: Dividing a PSU/cluster administrative area into equal-population segments when the area is too large to enumerate in the field. One segment is randomly selected to represent the PSU/cluster. Often, segmentation is performed in the field.

Mapping: The process of mapping all buildings, streets, and landmarks in a PSU/cluster.



Figure 3. Segmentation of a PSU/cluster into two approximately equal-population areas

Listing: The process of listing all dwellings (one-stage design) or households (two-stage design) in each building. In two-stage sampling, the head of household name is collected for each household.

Enumeration: This term is sometimes used interchangeably with “mapping”, or “mapping and listing”, or even “mapping, listing, and interviewing”. Thus, given its multiple meanings, we avoid the term in this document.

Dwelling: A physical housing unit, usually distinguished by a separate entrance, and occupied by a family, group of people living together or a single person. Definitions are similar across surveys (Table 1). Note that the definition of a dwelling is closely linked to the social concept of a household, and thus needs to be discussed with residents. This is particularly true in regions of the world where the word “dwelling” does not exist in the language.

Household: A social unit describing one or more people who normally sleep in the same dwelling, and who perhaps share economic resources or food. Definitions of a “household” vary considerably by survey (Table 2).

Table 1. Definitions of “dwelling” from major surveys

Survey	Definition
Demographic and Health Surveys	A room or a group of rooms normally intended as a residence for one household (for example: a single house, an apartment, a group of rooms in a house); a dwelling unit can also have more than one household.
Multiple Indicator Cluster Surveys	A room or a group of rooms normally intended as a place of residence for one household (e.g., a single house, an apartment, or a group of rooms in a house). However, a dwelling unit can also be shared by more than one household.
Living Standards Measurement Survey	A group of rooms or a single room occupied or intended for occupancy as separate living quarters by a family or some other group of persons living together, or by a person living alone.

Table 2. Definitions of "household" from major surveys

Survey	Definition
Demographic and Health Surveys	A person or a group of related or unrelated persons, who live together in the same dwelling unit, who acknowledge one adult male or female 15 years old or older as the head of the household, who share the same housekeeping arrangements, and are considered as one unit
Multiple Indicator Cluster Surveys	A person or a group of related or unrelated persons, who live together in the same dwelling unit, who share common living arrangements, who acknowledge the same person as the household head, who eat together and are considered as one unit.
Living Standards Measurement Survey	A group of people who share a roof and a cooking pot.

Typical household survey: workflow

Typical household surveys follow a two-stage sample design by first listing EAs from the last national census, then sampling EAs with PPS (Figure 4). In preparation for the second stage of sampling, a mapping-listing team is hired to visit each sampled EA. The team maps all buildings, streets, and landmarks, usually by hand on paper.

To ensure fieldwork is feasible, the mapper segments large EAs into two or more approximately equal-sized segments, randomly selects one segment, and maps that one segment. The lister makes a list of all dwelling units and households. Ideally the dwelling and household list is made by talking to residents and neighbours about household living arrangements, however, in practice listers often assume that 1 front door represents 1 dwelling, and that 1 household occupies each dwelling.

The mapping-listing protocols are usually not described in detail in survey documentation, so it can be difficult to discern the exact protocols used, and the quality of the mapping and listing data.

After the fieldwork, the survey planning team selects a sample of households from the household listing to be interviewed.

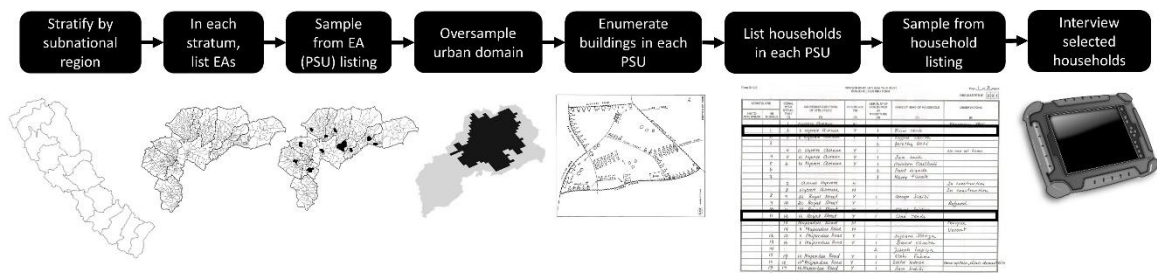


Figure 4. Typical household survey workflow

Typical household surveys: potential problems

Figure 5 outlines six ways that typical household survey workflow can exclude the urban poorest from being sampled, and two ways that the urban poorest are masked when data are collected. Each of these potential problems is described below.

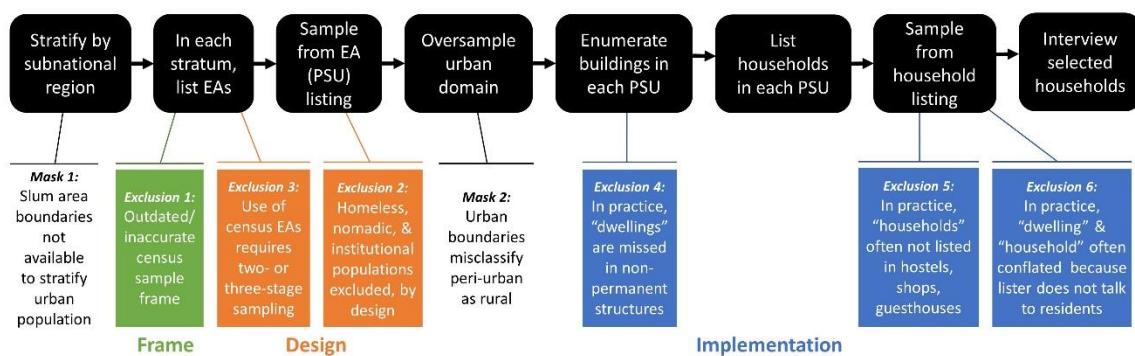


Figure 5. Problems with the typical household survey workflow

Exclusion 1: Outdated or inaccurate census sample frame. Urban populations are highly mobile and cities are growing rapidly, making administrative data from the last census outdated within a few years of it being collected. Nepal's last census was in 2011, and it is probably unreliable for sampling Kathmandu Valley in 2017 due to rapid population growth, and, possibly, population changes shaped by the 2015 earthquakes.

Exclusion 2: Homeless, nomadic, and institutional populations excluded from surveys, by design. While household surveys list all usual residents, surveys exclude "institutional" populations, the "homeless", and remote "nomadic" populations. Depending on how these terms are interpreted by the planning team and field

enumeration-listing staff, certain individuals and families who stay long-term in the city, participate in the city's economy, and use city services may be excluded. For example, people living in hostels and guesthouses may be classified as "institutional" and excluded in the same way that hospital patients and prisoners are excluded. People staying in migrant worker camps and construction camps may be interpreted as not being usual residents, or homeless.

Exclusion 3: Use of census EAs requires two- or three-stage sample. In the 1980s when survey methods were established, census EAs were generally the only population data available for sampling. The typical two-stage cluster sample design was established to (a) accommodate the large size of first-stage sample administrative areas, and (b) deal with similarities of households within sampling areas. However neither of these conditions necessarily hold true for cities today.

Today, other sources of population data such as gridded population datasets, are available for sampling, permitting sampling areas of different sizes and shapes. Furthermore, cities are highly mixed. Household living conditions and economic circumstances vary widely within the same neighbourhood, and sometimes, even within the same building. Thus it may not be necessary to assume similarity of households within sampling areas while designing the survey.

The use of EAs as the first-stage sample frame has other consequences which can lead to exclusion of the urban poorest. In typical survey enumeration, all buildings are mapped and listed months before the interview team conducts interviews with residents. This can be a problem in dynamic city contexts if the mapping and household listing become outdated during that time period. Survey implementers assume that dwellings vacated by one household will be replaced by a similar household (thus not changing the accuracy of the sample in terms of its representivity of the population). However, this assumption does not account for seasonal working patterns which might lead to different types of households being present during the mapping-listing period versus interview period a few months later.

Exclusion 4: In practice, dwellings are missed in non-permanent structures. Typical household surveys aim to document all usual residents, including slum dwellers. However, without clear training, mappers-listers might not include informal structures such as workers staying long-term in tents on a construction site, or shacks located behind a permanent building. People who live in non-permanent buildings almost always have low social-economic status, and thus may be under-listed and under-represented in the final sample.

Exclusion 5: In practice, households are routinely not listed in hostels, shops, and guesthouses. Household surveys also aim to include non-standard types of households including guards or cleaning staff who stay long-term in a commercial building, single adults who work in the city, and individuals or families who live at their business. If mappers-listers do not have detailed conversations with residents and neighbours about their living arrangements, including at non-residential buildings, then a range of households will be under-listed and under-sampled.

Exclusion 6: In practice, dwelling and household often conflated because lister does not talk to residents. Given the multitude of living arrangements in cities, and the increasing cost of living which forces lower-income families to share dwellings, listers cannot assume that one household occupies each dwelling. If these assumptions are made, poorer individuals and households will be under-listed and under-sampled.

Masking the poorest 1: Slum area boundaries are not available to stratify urban populations. There is no universal definition of a “slum area”, and numerous characteristics may define a slum area including unplanned buildings, lack of tenure, or lack of municipal-provided infrastructure. Slum area boundaries also change rapidly due to in-migration, government demolitions, and environmental catastrophes such as landslides or floods. Slum areas are highly likely to be missing from the sample frame, though even when slum areas are included in household surveys, the data collected in these neighbourhoods is averaged with data collected from all other households in the city, masking the unique health risks and outcomes experienced by this population.

Masking the poorest 2: Outdated urban boundaries misclassify peri-urban populations as rural. Sometimes urban boundaries do not include peri-urban areas. If peri-urban areas contain slum communities with residents who participate in the city’s daily economy and social life, then those populations should be counted with urban populations. By using a boundary definition which excludes peri-urban residents, then peri-urban residents get counted as rural, they do not get oversampled with urban populations, and their urban vulnerabilities get masked.

SUE surveys

Let us review a few more terms to discuss the SUE Survey workflow.

“Top-down” gridded population data: A model of the population that disaggregates census data to small grid squares. At the time of this writing, two “top-down” gridded population datasets are being used to select household survey samples: WorldPop (~100 metre by 100 metre grid squares, the size of a city block) and LandScan (~1 kilometre by 1 kilometre grid squares). Other “top-down” gridded population datasets are available, but are much less accurate per grid square (GPWv4, GRUMP, GHS-POP) or are only available in a few countries (Demobase, HRSL). To disaggregate census data, WorldPop and LandScan models use multiple spatial datasets that indicate where populations live (such as roads, rivers, buildings, land cover type, elevation, night-time lights, etcetera). LandScan models the “ambient” 24-hour population which is the average of night-time residential population and the day-time commuter population, and this dataset is not free. WorldPop models the night-time residential population, is well documented (if you want to understand the datasets used to model it), and is freely available. WorldPop is recommended over LandScan, particularly for selection of urban household surveys, however both datasets could be problematic in areas with grossly outdated census data.

“Bottom-up” gridded population datasets. At the time of this writing, no census-independent “bottom-up” gridded population datasets were available, but they could become available soon. “Bottom-up” gridded datasets are models of the population generated from high resolution satellite imagery and micro-census data. The accuracy of these datasets remains to be evaluated, however, they are promising in countries with grossly outdated or inaccurate census data.

WorldPop: WorldPop is a gridded population dataset that uses a complex geostatistical model to disaggregate census population counts to 100m X 100m grid cells (about the size of a city block). While the population totals may be inaccurate because they are based on potentially outdated census data, the distribution of the population may be more accurate than the census data because the modelled distribution of the population reflects the spatial datasets which include informal settlements and new construction. WorldPop datasets are available for over 140 countries for free at www.worldpop.org.

Spatial sampling: Oversampling in space such that we force there to be at least one PSU per large area (the survey designer defined the dimensions of this large area).

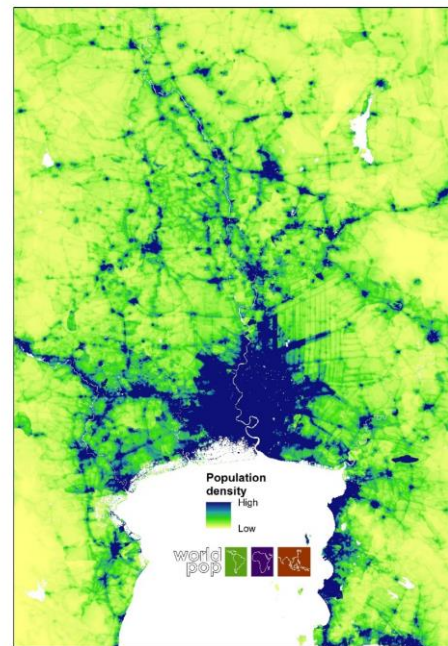


Figure 6. WorldPop gridded population data for Dhaka

GridSample: GridSample.org is a web-based tool for selecting representative household survey samples from gridded population data. WorldPop datasets, administrative boundaries, and urban/rural boundaries are preloaded on the website to specify a household sample design. Among other survey parameters, GridSample allows the user to specify the maximum geographic size or desired population size of sampling areas.

One-stage PSU: In the Nepal SUE survey, one-stage sampling areas are comprised of one grid cell. One-stage sampling areas in the Bangladesh and Vietnam SUE surveys may be comprised of multiple cells. The use of one cell PSUs in Nepal was the result of randomly assigning PSUs to the one-stage versus two-stage arm of the study.

Two-stage PSU: In the Nepal SUE survey, two-stage PSUs contain approximately 200 households each, and are usually comprised of multiple cells.

Digital segmentation: Digital segmentation is performed in the office using satellite imagery in a GIS. Segment boundaries follow roads, rivers, or other features that are easy to locate in the field.

Sample weight: The probability that a specific household/individual is included in the final survey. The sample weight adjusts for unequal, but known, probabilities of inclusion due to stratification, oversampling, segmentation, and non-response. Sample weights are generated after the survey, and are applied during survey data analysis.

OpenStreetMap (OSM): OpenStreetMap is a free, public, crowd-sourced world map of streets, buildings, landmarks, and many other features that are updated continually. www.openstreetmap.org.

iD Editor: iD Editor is OpenStreetMap's easy, embedded tool to edit features. iD Editor allows users to display a GPX file of boundaries (such as PSUs) on top of OpenStreetMap layer without sharing these data publicly.

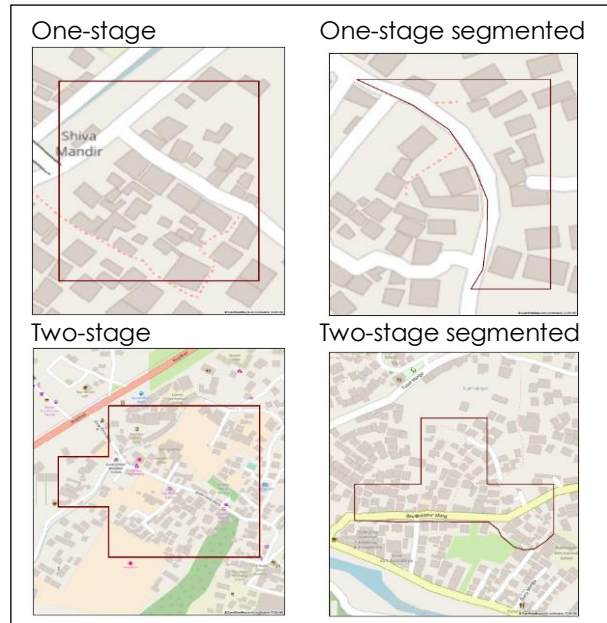


Figure 7. Visuals of gridded one-stage and two-stage sampling areas

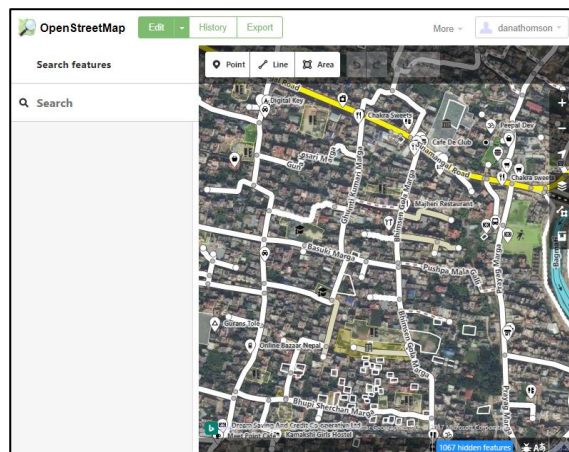


Figure 8. iD Editor to edit OpenStreetMap

OSMAnd: A mobile app that works offline and simultaneously displays OpenStreetMap data, satellite imagery, boundaries of PSUs, and the user's location.

Digital enumeration: This refers to mapping buildings, streets, and landmarks digitally in order to produce a geographically accurate map for fieldwork.

Digital listing: This refers to the use of a mobile device to list all dwellings and households in selected PSUs.

GeoODK: GeoODK is an Open Data Kit application that works offline on a tablet. It has the capacity to collect geographic location and photographic information with each entry.

Geographic information system (GIS): A software used to generate, store, manage, modify, analyse, and present spatial data. The two most common GIS tools are ArcGIS, a very common commercial product, and QGIS, a free, open-source product. The SUE manuals provide instructions for ArcGIS, though all steps can be performed in QGIS.

Permanent building: Permanent buildings may have one or more dwellings, and may be residential, commercial, a mix of commercial and residential, or institutional. The SUE survey considers any building with brick or cement walls to be permanent, and will be updated in OpenStreetMap.

Non-permanent building: The SUE surveys consider non-permanent buildings to include shacks (made of durable materials such as a tin roof or wood panel walls)

and tents (made of non-durable materials such as a tarp roof or cloth walls). In other words, the SUE surveys consider a building non-permanent when the walls are *not* made of brick or cement. Residential shacks and tents are not entered in OpenStreetMap, and are instead digitally recorded offline in ArcGIS. Note, this definition of a "non-permanent building" might need to be updated for other contexts, for example, if wood is used to construct permanent dwellings.



Figure 9. Non-permanent buildings

Level (of building): The SUE surveys record the number of levels in each building, and link each dwelling to one level. A level will be counted if it has walls and a roof. This means that covered porches, stair shelters and storage sheds on the top of buildings should not be counted as levels. Levels are always numbered from bottom-to-top, starting with level 1.

Dwelling: The SUE surveys use a nuanced definition of dwelling: a covered physical housing unit comprised of 1+ rooms where people sleep; often distinguished by a separate entrance, however you may step outside to enter different rooms; and occupied by 1+ families, groups of people living together, or single persons. If there are multiple dwellings on the same level, list dwellings from left-to-right from the perspective of the building/compound main entrance.

Household: The Nepal two-stage SUE survey uses the Demographic and Health Survey definition of a household during listing activities. Listers describe a household as “a group of people who normally share a cookpot.”



Figure 10. Listing levels of a building

SUE Survey: Workflow

The SUE surveys follow the same workflow as typical household surveys, however the SUE surveys use different tools and methods. In Nepal, half of the survey is a one-stage sample for which a list of dwellings is prepared, and half of the survey is a two-stage sample for which a list of households is prepared.

GridSample is used to select PSUs, comprised of one or more cells, from WorldPop data. In selected PSUs, all buildings and streets are digitally mapped. First, buildings and streets are enumerated in OpenStreetMap at the office based on satellite imagery using iD Editor. Second, updates to building locations and levels are made in the field on printed OpenStreetMap maps using OSMAnd as a reference, and this information is entered in OpenStreetMap via iD Editor in a third step. Note, shack/tent-residential buildings are not entered in OpenStreetMap, and instead are entered in ArcGIS offline in a later step.

In the one-stage PSUs, listers list dwellings in GeoODK based on conversations with residents and neighbours. In two-stage PSUs, listers list households in GoeODK, including head of household name, based on conversations with residents and neighbours. GeoODK data is synced to a server daily.

To prepare for one-stage survey interviewers, the SUE survey planning team further segments one-stage PSUs until they contain 15-20 dwellings each. To prepare for the two-stage survey interviews (in Nepal only), the planning team randomly samples 20 households in each two-stage PSU.

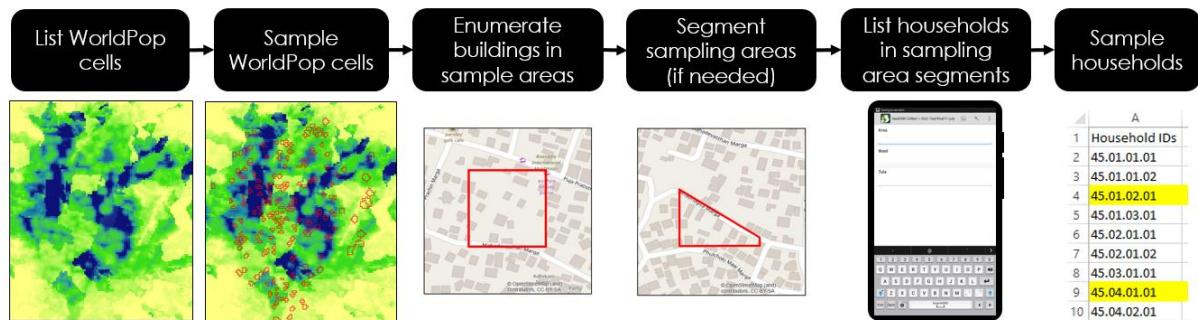


Figure 11. SUE Survey Workflow (one-stage survey)

SUE surveys: addressing potential problems

In all SUE surveys, 2017 WorldPop data is **expected to reflect areas of new population growth and unregistered slums, and thus be more accurate of the population distribution** than census data. GridSample allows selection of PSUs of a specified size – 20 households for the one-stage surveys, and 200 households for the two-stage survey.

We considered use of municipality boundaries to define the urban extent of Kathmandu Valley, however, “metropolitan” boundaries only covered the downtown area of Kathmandu, and adding “sub-metropolitan” boundaries would have resulted in a substantial sample of rural communities beyond the peri-urban reach. Instead, we defined the Kathmandu Valley boundary using the **GHS-SMOD model of urban extents which accounts for both built-up landcover and population density** associated with the city.

The use of OpenStreetMap for building enumerations **leverages existing data**, and where that data is lacking, **contributes back to a public resource**. ID Editor allows some of this enumeration work to be completed in the comfort of the office before and during fieldwork, **minimizing time in the field**. Furthermore, the use of printed paper OpenStreetMap maps in the field permits quick note taking (versus making digital edits in the field on an app such as Vespucci), and allows for **shack/tent-residential buildings to be recorded offline privately**.

The use of GeoODK allows for collection of a **GPS coordinate and a photograph for each building**, and this data is stored securely in a database. A core feature of the SUE methodology is a detailed script and protocol to converse with residents and

neighbours during the listing process to collect an accurate, complete list of dwellings and households.

SUE surveys: evaluation

The SUE survey two-stage sample provides sufficient data (600 households) to compare demographics of respondents to respondents in the 2016 Demographic and Health Survey. This will allow us to **assess the extent to which SUE survey methods and tools result in a different population sample compared to typical household surveys** (for example, whether SUE survey methods and tools resulted in a larger proportion of poor households).

The SUE survey one-stage sample in Nepal (600 households) will include additional types of populations not included in typical household surveys, and will help **to assess whether these additional household types should be included in future household surveys**.

By comparing the demographic characteristics of the SUE survey one-stage sample and two-stage sample, we will **gather evidence about whether the additional time and effort involved with two-stage sampling is necessary to collect population-representative data in a heterogeneous complex urban environment**.

The one-stage samples in Bangladesh and Vietnam (400 households each) will provide sub-city estimates of key indicators for local decision-makers and will evaluate whether the SUE survey tools and methods developed in Nepal are scalable across multiple different complex settings.

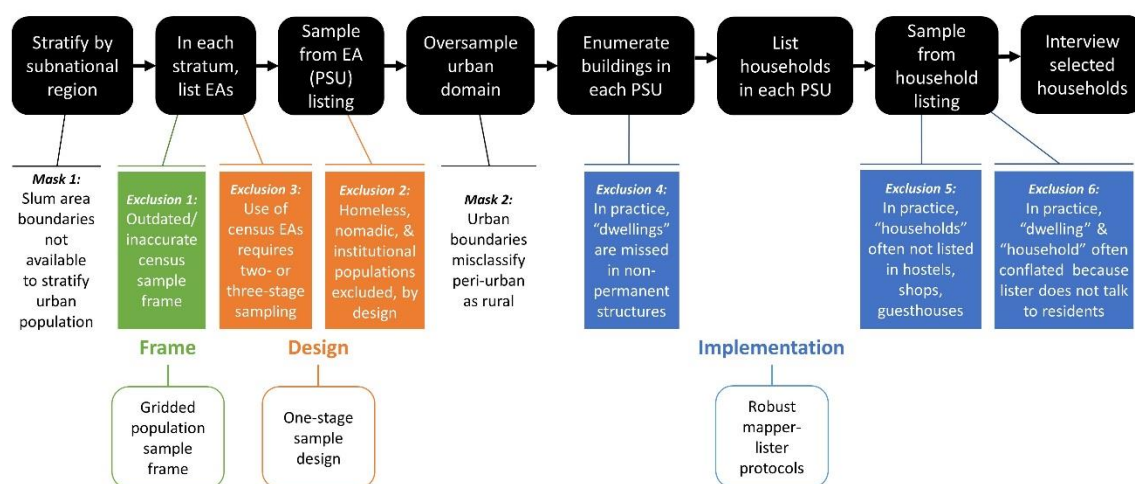


Figure 12. Addressing potential problems in typical household survey workflow

Our initial experience with SUE survey methods identified several challenges to implement these methods in complex urban environments which required creative problem-solving. Table 3 summarizes the challenges we faced, and our actions to overcome them.

Table 3. Challenges and suggestions to overcome limitations of SUE survey methods

Challenges	Overcoming challenges
Free imagery may have low resolution	<ul style="list-style-type: none"> • Use multiple tools with imagery including Bing in OpenStreetMap, Google Earth, and Google Maps • If you can, budget to purchase high resolution imagery from a commercial company
Rapid urbanization means rapid change in building and street infrastructure; Imagery not updated in timely manner	<ul style="list-style-type: none"> • Use multiple tools with imagery including Bing in OpenStreetMap, Google Earth, and Google Maps • Involve people with local place knowledge in the pre-field segmentation process. Perform fieldwork in areas where populations are thought to live but do not appear in older imagery
Limited or "patchy" mobile phone coverage	<ul style="list-style-type: none"> • Use only tools that function offline
Multiple applications need to be used, and some apps require high level of technical expertise to set up	<ul style="list-style-type: none"> • Schedule sufficient time to learn new methods and tools. • Identify technical experts to be involved. • Looking forward, we recommend that software developers integrate the functionality of separate tools into one tablet-based survey toolbox
Mapper-lister staff may be different than typical household surveys	<ul style="list-style-type: none"> • Advertise Mapper-Lister-GIS position in university geography or geomatics engineering departments rather than public health departments
OpenStreetMap (OSM) updates sync in different time to ArcGIS, GeoODK, OSMAnd	<ul style="list-style-type: none"> • In our experience, ArcGIS syncs OpenStreetMap base map almost immediately. • OSMAnd should sync OpenStreetMap quickly if using the "Pro" paid version with a low-cost monthly subscription. • It is unclear how long it takes for GeoODK to update OpenStreetMap.
Costs and time	<ul style="list-style-type: none"> • We are still evaluating how the time and costs of these methods compare to typical household surveys.

Review

Answer the questions below. To see if your answers are correct, go to Appendix D.

1. WorldPop is a:
 - A. Aggregated global census dataset of the population
 - B. Gridded population model of the population
 - C. A tablet-based app for estimating population totals in areas
 - D. A toolkit in ArcGIS using during segmentation

2. The definition of a "dwelling" use by SUE surveys is:
 - A. A physical housing unit
 - B. Distinguished by a separate entrance
 - C. Occupied by a family, group of people living together, or single person
 - D. All of the above

3. In SUE surveys, list dwellings in this order when you are looking at the front of the building (or main entrance)
 - A. Bottom-to-top
 - B. Left-to-right
 - C. Top-to-bottom
 - D. Right-to-left

4. Sampling area segments are:
 - A. Approximately equal geographic size
 - B. Approximately equal population size
 - C. Always half the original sampling area size
 - D. Always 1/3 the original sampling area size

5. In two-stage PSUs in the SUE surveys, four pieces of critical information are collected about each household by the mapper-lister team. What are these four pieces of information?

3. Workflow

Objectives:

By the end of this chapter, you will be familiar with:

- The key tasks performed, by whom, where, and timing throughout the survey
- The tools used to perform each task
- How one-stage and two-stage samples are similar and different

Outline:

- Overview
- Survey design
- Sample selection
- Enumeration and listing
- Interviewing

Overview

Figures 12 and 13 summarize the Nepal, Bangladesh, and Vietnam SUE survey workflows. Tasks are performed by **four types of staff**: Planning Team members, Mappers-Listers, GIS staff, and Interviewers. Staff perform either office-based or field-based work, indicated with different symbols.

In Nepal, the Planning Team is comprised of HERD International staff Radheshyam Bhattarai, Rajeev Dhungel, and Subash Gajurel, plus Dana Thomson from University of Southampton. The Planning Team in Bangladesh includes Silvia Ferdoush from CIPRB and Tarana Ferdoush from ARK Foundation. In Vietnam, the Planning Team includes Duong Minh Duc and Nguyen Bao Ngoc from the Hanoi School of Public Health. The Planning Teams all work closely with the SUE project coordinated by Helen Eley at Leeds University in the UK.

In Kathmandu, enumerator-lister staff are recruited from Kathmandu University Geomatics Engineering Department. The GIS team is comprised of one Planning Team member, Dana Thomson, plus a HERD International GIS Intern. Interview staff are recruited from local public health university programmes.

Six tools are used throughout the SUE survey workflow: GridSample, iD Editor, ArcGIS, OSMAnd, GeoODK, and Excel. Each will be described in the next section.

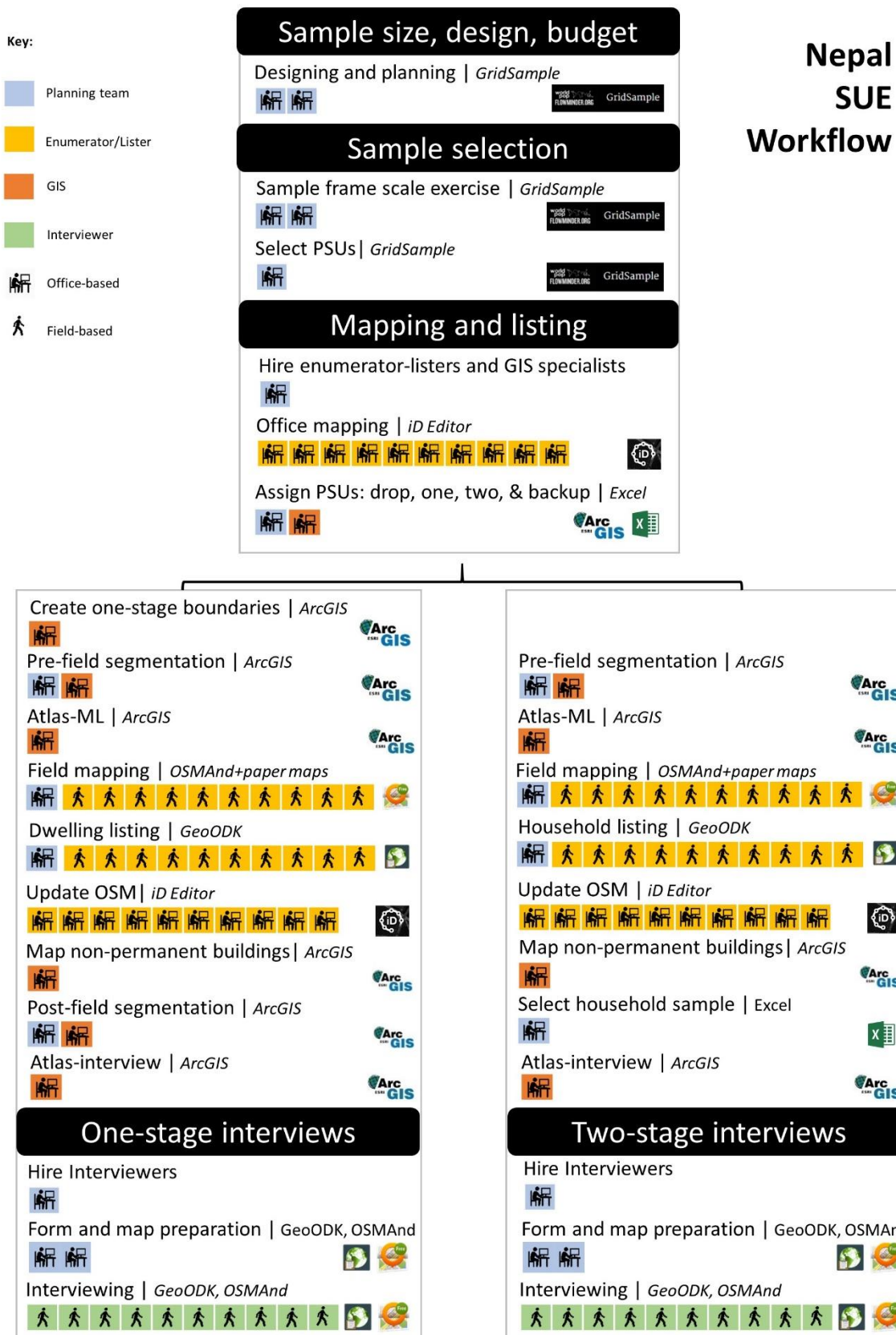


Figure 13. SUE Survey Workflow in Nepal

Key:

- Planning team
- Enumerator/Listener
- GIS
- Interviewer
- Office-based
- Field-based

Bangladesh & Vietnam SUE workflow

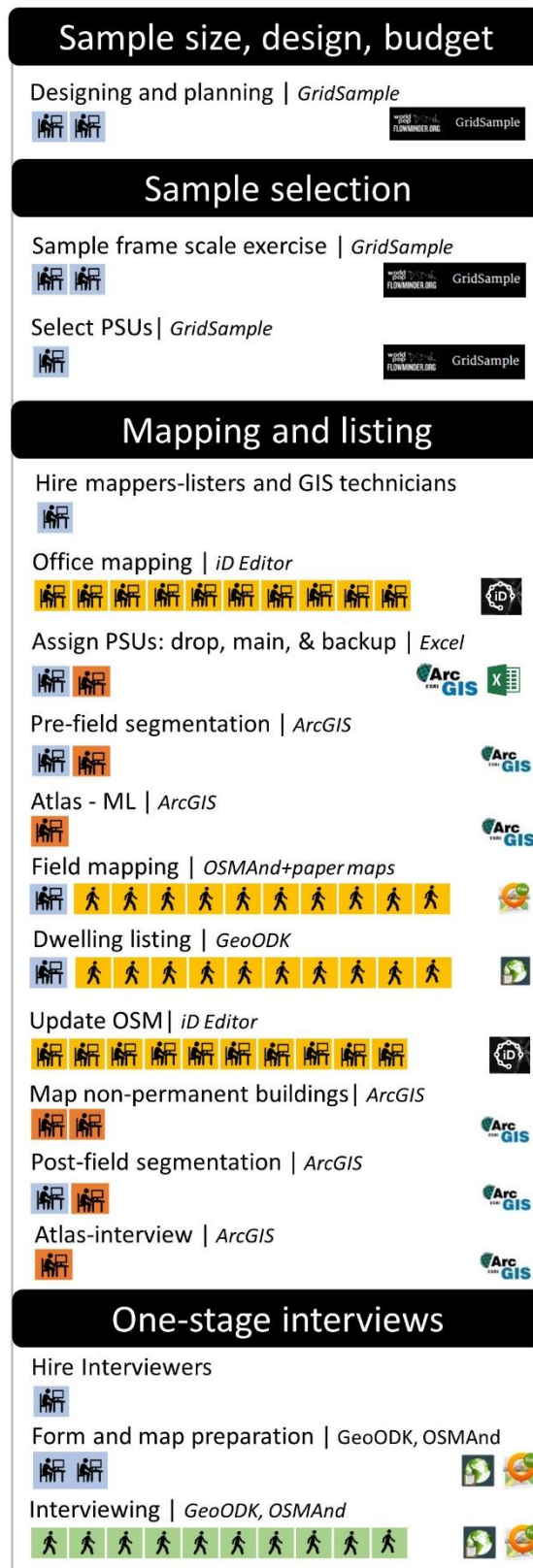


Figure 14. SUE Survey Workflow in Bangladesh and Vietnam

Survey design

The Nepal SUE survey has 60 PSUs across the Kathmandu Valley – 30 PSUs are randomly assigned to be one-stage samples, and 30 PSUs to two-stage samples (see Chapter 7. Logistics: Pre-field for details). In each PSU, approximately 20 households are interviewed for a total of 1200 households. An additional 18 PSUs (30% extra) are selected as back-up sampling areas. Back-up sampling areas are needed if the PSUs have little-to-no population because, for example, it is located on the airport runway.

The Bangladesh SUE survey has 15 one-stage PSUs (plus 5 back-up PSUs) in Ward 56 of Kamrangichar thana, South Municipality, and 5 one-stage PSUs (plus 3 back-up PSUs) in Bhola slum. 15-20 dwellings are sampled in each PSU, for a total of 400 interviews.

The Vietnam SUE survey has 20 one-stage PSUs (plus 10 back-up PSUs) in Long Bien district. 15-20 dwellings are sampled in each sampling area for a total of 400 interviews.

Mapping and listing

Mapping is comprised of three steps:

- (1) Office mapping in OpenStreetmap
- (2) Field mapping on printed paper maps with OpenStreetMap features
- (3) Updating OpenStreetMap with field notes

Listing dwellings and households involves detail conversations with residents and neighbours of each building. One-stage PSUs are geographically small. The listing is of approximately 15-20 dwellings per PSU – though possibly more in dense downtown locations since we cannot distinguish residential and other buildings types from satellite imagery before fieldwork. Two-stage PSUs are much larger. The listing includes approximately 200 households, including the name of each head of household.

There are a number of **GIS** processing steps including:

- Atlas map generation
 - Version ML: Before field enumeration with the segmented sampling areas
 - Version Interview: After field enumeration and after one-stage sampling areas have been segmented a 2nd time – these are the final maps that the interview team will use

- Creating single cell boundaries for the one-stage sample (Nepal only)
- Segmentation of large sampling areas
- Mapping shack/tent-residential buildings offline in ArcGIS

Interviewing

For both one and two-stage, all participants must be over 18 and in a fit mental state and not under the influence of any substance.

In two-stage PSUs, interviewers return to 20 randomly selected households and administer a household questionnaire, plus an individual questionnaire to one randomly selected adult. The household questionnaire includes detailed questions about household wealth and injuries experienced by every member of the households. The individual questionnaire includes questions about mental health and wellbeing, and opinions of the selected adult. In a two-stage PSUs, servants, non-relatives, and tenants who share the dwelling are recorded as part of the main household.

In one-stage PSUs, the interviewer returns to 15-20 dwellings located next to each other and identifies all households to administer the household and individual questionnaires described above. In one-stage sampling the aim is to include all urban residents who might not be represented in traditional two-stage designs. This means interviewing **individuals or families who are living with others in a guest house or hostel, those sleeping on the street and those who are workers or non-relatives** living in a dwelling with another household. In one-stage PSUs, record servants, non-relatives, and tenants (including tenant's relatives) as separate households.

Who to interview in hostels and guesthouses:

- If a **hostel is only for education or health-care purposes** - such as a student hostel, boarding school, or an old-age home - then **all of the residents do NOT need to be interviewed**. Instead conduct the interview with any resident warden or supervisor and make a count of the number of people living there. This household would be classified as an institution or college dorm.
- If a **hostel or guesthouse** has people **living there for more than 7 days** who are working, looking for work, or in the city for any other purpose (such as supporting someone in hospital or other institution separate from the hostel/guesthouse), then they **should be interviewed as individuals** such that each complete their own questionnaire. If their children or other relatives are also staying with this individual then they can complete the questionnaire as one household. This would be classified as residential.

Interviewing Street Sleepers:

The one-stage sampling allows us to conduct the questionnaire with any street-sleepers that fit our eligibility criteria (over 18 years and in a fit mental state meaning they are not under the influence of any substance or severely mentally ill).

Identifying street sleepers to interview is going to take special skills and timing of the interview team. This may require interviewers to go early in the morning, and to go in pairs (and to be male if needed) for safety. It will also require sensitivity in conducting the questionnaire, and possible modification of some questions, as many of the questions may not be appropriate (e.g. housing materials).

Table 4. Comparison of one-stage versus two-stage workflows in SUE Surveys

Description	One-stage	Two-stage
Sample size	30 sampling areas	30 sampling areas
Unit of final listing	Dwelling	Household
Size of PSU	15-20 dwellings	150-200 households
Difference in sampling process	Post-field segmentation in one-stage sampling areas, not two-stage sampling areas	
Mapping	Map all buildings, streets, and key points of interest	
Listing	Residents, including hostels Long-term guards, staff, workers Long-term squatters	
Excluded from listing	Residents of embassy compounds, prisoners, army barracks, school dorms, hospital patients, street-sleepers, long-term guesthouse residents	
Respondents to interviews (Over age 18, fit mental state, not under the influence of substances)	Residents, incl. hostels Guards, staff, workers Long-term squatters Guesthouse guests who stayed for the last week Street-sleepers	Residents, incl. hostels Guards, staff, workers Long-term squatters
Excluded from interviews (Age 18 or younger, not in a fit mental state, under the influence of substances)	Residents of embassy compounds, army barracks, high school dorms, prisoners, hospital patients	Residents of embassy compounds, prisoners, army barracks, school dorms, hospital patients, street-sleepers, guesthouse residents

Review

Answer the question below. To see if your answer is correct, go to Appendix D.

1. The three main differences between the one-stage and two-stage samples is:

- a. Sampling process, mapping methods, listing methods
- b. Sample size, size of PSU, respondents to interviews
- c. Unit of final listing, size of sampling area, respondents to interviews
- d. Unit of final listing, mapping methods, listing methods

2. The main steps of the SUE mapping-listing team are:

- a. Map every resident and visitor on day of fieldwork, find 20 heads of households at home
- b. Office mapping, field mapping and listing, interviewing
- c. Pre-field segmentation, field segmentation, post-field segmentation
- d. Office mapping, field mapping and listing, post-field updates to OpenStreetMap

3. Which tool is used for updating OpenStreetMap

- a. ArcGIS
- b. iD Editor
- c. GeoODK
- d. OSMAnd

4. Survey Design and Selection

Objectives:

By the end of this chapter, you will be able to:

- Select a sample design and sample size that meets your budget
- Determine an appropriate gridded sample frame scale
- Use GridSample to select primary sampling units (PSUs)

Outline:

- Sample size, design, budget
- GridSample
- Sample frame scale
- Select PSUs

Sample size, design, budget

Before using GridSample to select sampling areas, the user must have a clear research question, decide on the survey design, and calculate the sample size per the available budget.

Identify primary question and acceptable level of uncertainty. All three of the SUE surveys will test the feasibility of novel survey methods including one-stage sampling. To evaluate whether these new methods are more likely than traditional (DHS-style) methods, to include the urban poor, the Nepal survey is larger than the Bangladesh and Vietnam surveys, and includes two-stage sampling areas which can be compared to the 2016 Nepal DHS.

Survey sample design generally fall into three categories: classification, estimation, or comparison. Table 3 summarizes the three survey design options as described in the WHO Vaccination Survey Coverage Guidelines, and on the [GridSample website](#).

Classification questions label areas as “pass”, “marginal”, or “fail” with respect to a predefined threshold to inform programmatic decisions. For example, a classification survey could be used to identify wards with high incidence of pedestrian and traffic-related injuries to help a municipal government plan and prioritize public safety projects. Most household surveys are designed to answer an estimation question such that statistical estimates (e.g. prevalence rate) can be estimated in one or more groups. Finally, a comparison survey makes a quantitative estimate of difference between two groups, or of change in one group between two time points.

Table 5. Summary of research questions and survey design options (adopted from WHO Vaccination Coverage Survey Manual, and reprinted from www.gridssample.org with permission)

Classification question	Estimation question	Comparison question
A classification survey labels groups as “pass”, “marginal”, or “fail” to inform programmatic decisions	An estimation survey results in quantitative estimates in one or more groups	A comparative survey make a quantitative estimate of difference between two groups, or of change between two time points
Example question: Which wards have high incidence of traffic-related injury?	Example question: What is incidence of injury by ward?	Example question: Has the incidence of injury in each ward changed since the last survey?
<p><i>Inferential goal:</i></p> <ul style="list-style-type: none"> • Pass/fail threshold <p><i>Uncertainty is reported as:</i></p> <ul style="list-style-type: none"> • Misclassification of pass • Misclassification of fail 	<p><i>Inferential goal:</i></p> <ul style="list-style-type: none"> • Coverage estimate <p><i>Uncertainty is reported as:</i></p> <ul style="list-style-type: none"> • Confidence interval <p><i>Design parameter:</i></p> <ul style="list-style-type: none"> • α (alpha) - the probability of Type 1 error - the hypothesis test declares the difference to be statistically significant when, in truth, it is not 	<p><i>Inferential goal:</i></p> <ul style="list-style-type: none"> • Minimum detectable difference between 2 groups <p><i>Uncertainty is reported as:</i></p> <ul style="list-style-type: none"> • Confidence interval <p><i>Design parameter:</i></p> <ul style="list-style-type: none"> • α (alpha) - the probability of Type 1 error - the hypothesis test declares the difference to be statistically significant when, in truth, it is not • $1-\beta$ (beta) - the power of the test - this is the probability that the hypothesis correctly identified a statistically significant difference
<p>Example parameters:</p> <ul style="list-style-type: none"> • Pass/fail threshold: 90% • Misclassify passes: 5% • Misclassify fails: 10% 	<p>Example parameters:</p> <ul style="list-style-type: none"> • $\alpha = 95\%$ (95% confidence interval) 	<p>Example parameters:</p> <ul style="list-style-type: none"> • difference = 10% or more • 2-sided hypothesis test • $\alpha = 95\%$ (95% confidence interval) • power = 80% ($\beta = 20\%$)
Generally requires 15+ PSUs per stratum	Generally requires 30+ PSUs per stratum	Sample size is highly dependent on the difference between groups and design parameters. 30-60 PSUs per stratum.

The **Nepal SUE survey** was designed as an estimation survey, meaning that the survey results estimates prevalence of key indicators among residents of the Kathmandu Valley. The sample had 60 sampling areas (30 one-stage sampling areas, and 30 two-stage sampling areas) of 20 households each, for a total of 1200 households. This is based on a prevalence of 50% +/- 4.27% margin of error with 95% confidence for a clustered survey with 1.41 mean design effect and household (98%) and individual (93%) response rates from the 2011 Nepal DHS. Refer to Appendix A for exercise to reproduce this sample size calculation.

The **Vietnam and Bangladesh SUE surveys** were designed to evaluate feasibility of methods in two local areas of Hanoi and Dhaka where complex urban conditions were known to exist including informal settlements. The one-stage sampling design of 20 households in 20 PSUs may be too small to make direct estimates of many key indicators, but can be used to classify survey areas as "low", "marginal" or "high" on those indicators. For example, this sample size will identify whether the area is higher than 70%, marginal (63-70%), or definitively lower than 70% (<63%) with 95% confidence and 80% power.

Generate multiple survey scenarios. Calculating a sample size that both meets the inferential goals and the budget constraints is an iterative process. The sample size + budget spreadsheet available on the [GridSample website](#) allows the user to generate multiple side-by-side scenarios to judge the feasibility of different sample sizes and expected costs. A budget for the Nepal survey is provided below (Table 5).

Table 6. Approximate budget for Nepal SUE survey

Item	Cost (USD)
Research and development	
Salaries	4,300
Consultant	1,400
Mapping-Listing-GIS	
Salaries, per diem	19,000
Materials (e.g. tablets)	4,800
Interviews & data management	
Salaries, per diem	23,800
Materials (including piloting questionnaire)	3,100
Ethics review	2,000
Overhead	
20% of direct costs	11,680
TOTAL	70,080

GridSample

The GridSample website (available at www.gridsample.org) guides the user through a sequence of steps to choose pre-loaded datasets and parameters, and generate PSUs for a cluster survey. In addition to specifying typical survey parameters (such as coverage area, and number of households and PSUs per stratum), the user may also set several gridded population sampling parameters: the size of the grid cells in the sample frame (sample frame scale), and how to define PSUs. Explore the website for further details.

Here is a checklist to determine whether you are ready to use GridSample:

- ✓ I know the coverage area of the survey
- ✓ I know the year the survey will be implemented
- ✓ I know whether the survey will be stratified, and the geographic boundaries of those strata
- ✓ I know the sample size per strata (number of PSUs, and households per PSU)
- ✓ I know whether the sample will be a one-stage or two-stage sample, and the total target population per PSU
- ✓ I know whether low density areas will be oversampled to ensure a population+spatial sample (not applicable for most household surveys)

Sample frame scale

The appropriate sample frame scale (size of grid cells in the gridded population dataset) depends on:

- The **accuracy of the gridded population sample frame** – A sample frame generated from outdated, less-detailed census data and covariates will be less accurate at fine geographic scale than a sample frame generated from recent, detailed census data and covariates. Aggregating the sample frame to a coarser scale (larger grid cells) usually improves the accuracy of the sample frame. Aggregating the sample frame too much, however, reduces the feasibility of fieldwork. Identifying the ideal scale of sample frames to balance feasibility and accuracy is an active area of research, and unfortunately limited guidance currently exists.
- The **field methods** to be used – the Planning Team may deliberately generate sampling areas with more population than will be visited during field enumeration in order to manually segment PSUs before fieldwork, creating segments with “sensible” boundaries which follow rivers, streets, and other geographic features that are easy to identify on the ground. Although this adds an extra GIS processing step to the workflow, it can improve survey feasibility by ensuring the PSU segment boundaries are easy to identify in the field.

In the absence of guidance about the ideal scale of the gridded population sample frame, we recommend completing the Sample Frame Scale exercise in Appendix A. This exercise guides the user to compare (a) the estimated population in the WorldPop sample frame, and (b) the number and type of buildings visible in satellite imagery. The ideal sample frame has minimal discrepancies between these datasets, and has sample frame and cells that are close in geographic size to the final PSUs.

Defining PSUs

None of the SUE Surveys were selected via the GridSample.org because at the time of this work the GridSample Beta website did not allow users to define custom coverage area boundaries. Instead, the GridSample R algorithm was used. Below are a summary of the parameters used in the GridSample R algorithm, along with a visual of the customized coverage area boundaries.

For the Nepal SUE survey, we selected a two-stage sample in which PSUs had approximately 200+ households each. To generate the one-stage PSUs in Nepal, we selected the centroid 100m X 100m cell of each sampling area. This meant that one-stage PSUs in Kathmandu are always a single cell. However, in Bangladesh and Hanoi, one-stage PSUs could be comprised of multiple cells.

Nepal SUE Survey

- Coverage: urban_ghssmod_wgs84.shp [custom dataset]
- WorldPop Dataset: NPL_ppp_v2c_2017_UNadj
- Stages: Two-stage
- Stratification: No
- Urban/rural domain: No
- Spatial representation: No
- Target population: "Household"
- Number of target population members per household: 1
- Average household size: 4.1
- Min pop/PSU: 820 population per PSU [~200 households per PSU]
- Number of PSUs per stratum: 78 [60 PSUs + 30% back-up]
- Number of households sampled per urban PSU: 20
- Number of households sampled per stratum: 1560 [78 PSUs * 20 HHs each]
- Input grid cells are: 100m X 100m [based on scale exercise]
- Exclude cells fewer than: 0.1 population
- Do not generate PSUs larger than: Infinity
- Random number: 2017
- Growth algorithm: on

Vietnam SUE Survey:

- Coverage: Long_Bien_district.shp [custom dataset]
- Dataset: VNM_ppp_v2c_2020_UNadj
- Stages: One-stage
- Stratification: no
- Urban/rural domain: no
- Spatial representation: no
- Target population: "Household"
- Number of target population members per household: 1
- Number of PSUs per stratum: 30 [20 PSUs + 50% backup]
- Number of households per PSU: ~20
- Number of households sampled per stratum: ~600 [30 PSUs * 20 HHs each]
- Input grid cells are: 200m X 200m [based on scale exercise]
- Exclude cells fewer than: 0.1
- Do not generate PSUs larger than: Inf
- Random number: 2017
- Growth algorithm: on

Bangladesh SUE Survey:

- Coverage: Ward56_Bhola.shp [custom dataset]
- Dataset: BNG_ppp_v2b_2020_UNadj
- Stages: One-stages
- Stratification: no
- Urban/rural domain: no
- Spatial representation: no
- Target population: "Household"
- Number of target population members per household: 1
- Number of PSUs per stratum: 20 [15 PSUs + 30% back-up]
- Number of households per PSU: ~20
- Number of households sampled per stratum: ~400 [20 PSUs * 20 HHs each]
- Input grid cells are: 100m X 100m [based on scale exercise]
- Exclude cells fewer than: 0.1
- Do not generate PSUs larger than: Inf
- Random number: 2017
- Growth algorithm: on

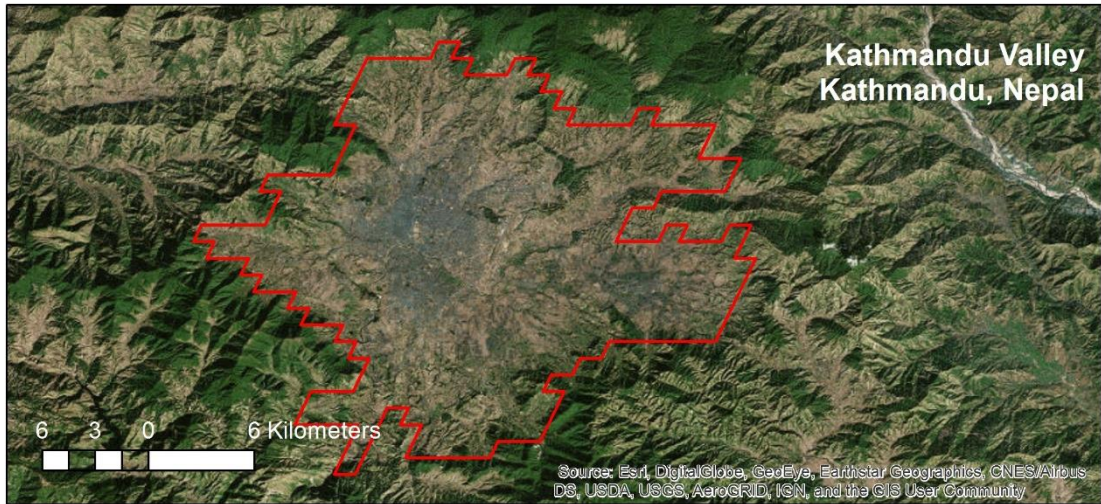


Figure 15. SUE survey coverage areas across all three countries

Review

Answer the questions below. To see if your answer is correct, go to Appendix D.

1. You are designing a household survey to measure the percent of individuals (adults and children) in a population who are uninsured. What type of research question is this?

- a. Classification question
- b. Estimation question
- c. Comparison question

2. You have completed the sample frame scale activity in Appendix A for 100m X 100m grid cells. You believe that the estimated population in WorldPop cells is far too high in more than half of the cells you compared to Google Earth imagery. What does this mean, and what should you do next?

- a. WorldPop > Google Earth is expected and you should move forward with final sample selection
- b. WorldPop > Google Earth is not expected, and you should not use gridded population sampling in your survey
- c. WorldPop > Google Earth is not expected, and you should aggregate grid cells to 200m X 200m (or possibly larger, e.g. 300m X 300m) and redo the sample frame scale activity until you identify a scale at which WorldPop = Google Earth.

3. "Sensible" PSUs have natural boundaries such as rivers and streets. How can you create "sensible" PSU boundaries?

- a. Use GridSample to select PSUs larger than needed, manually segment PSUs into areas of approximately equal population, then randomly select one segment
- b. All GridSample PSU boundaries are already "sensible".
- c. Use GridSample to select PSU centroids, then manually delineate PSUs around nearest [X number of] buildings visible in Google Earth

5. Infrastructure & Tools

Objectives:

By the end of this chapter, you will be familiar with:

- Infrastructure to support survey administration
- Tools used by Mapper-Lister Team
- Tools used by the GIS Team

Outline:

- Server
- Set-up of tablets, computers, and forms
- GIS workstation
- Tablets

Server

In the SUE surveys, the dwelling/household listing form and all questionnaires were administered on a tablet using the GeoODK app built on the Open Data Kit platform. The Planning Team may need support from an IT department to set-up a server to store survey forms, and manage submitted form data.

Please refer to “How To: Setup GeoODK server” in Appendix C for details about the HERD International server set-up. This document describes the HERD International server hosted at SMAP consulting, including its core functionality. Other teams should use any server system that is already familiar, as long as that system performs the same functions.

Set-up of tablets, computers, and forms

Tablets. Each mapper-lister was assigned a tablet for field work. Set-up of tablets can be performed by the Planning Team or an IT department. Detailed instructions are provided in Appendix C in “How To: Setup tablets” to install and configure each of the applications needed by mappers-listers. These instructions include steps to restrict access to only those applications used for the survey.

The instructions cover the following topics:

Update settings:

- Date and time
- Prediction (auto-complete) off
- Allow third party applications

Install applications:

- WiFi manager
- App Lock
- MAPS.ME
- Vespucci OSM Editor (Optional)
- GeoODK Collect
- Maps & GPS Navigation (OSMAAnd)
- Nova Launcher
- CamScanner-phone PDF Creator

Configure applications:

- App lock
- Nova launcher
- GeoODK - platform setup
- GeoODK - load XLS form
- GeoODK - restriction
- GeoODK - offline layer
- CamScanner
- MAPS.ME
- OSMAAnd

Forms. The GeoODK mapper-listing form can be designed and likely by the Planning Team. The listing form is programmed in Excel and then uploaded to the server. The server translates the Excel fields into questions and skip logic to be displayed in the tablet form. A survey template file is available to streamline the form set-up. See “How To: Setup GeoODK form in Excel” in Appendix C for instructions.


Computers. During office enumeration, each mapper-lister will use a computer with WiFi connection to update OpenStreetMap in iD Editor. No special software is needed. Having multiple mappers-listers simultaneously logged on to the internet will strain normal bandwidth allowances. To ensure efficiency and accuracy in their work, ensure usable computer work stations and sufficient WiFi bandwidth, specifically:

- Computer screens have decent colour and resolution
- Each computer is equipped with a mouse
- The WiFi bandwidth to computers is increased, if needed. In Nepal, the required bandwidth to update OpenStreetMap using 14 individual workstation was 2 to 3.5 MBPs.

GIS Workstation

At least one workstation should be set-up for the GIS Team. Table 6 summarizes the software that should be installed on this workstation, including alternative software that may be used.

Table 7. Software installed on GIS computer(s)






Purpose: App	Description	Alternatives considered
<p>GIS: ArcGIS</p> 	<p>The ArcGIS tool, ArcMap, is used to make maps, manipulate or create spatial data (segmentation, mapping shacks/tents offline), and perform spatial analysis. Use of ArcGIS requires intermediate GIS training. More information is available at www.arcgis.com.</p>	<p>QGIS (www.qgis.org) is a good and free alternative to ArcMap, and requires intermediate GIS training.</p> <p>Google Earth (www.google.com/earth/) is a much less technical, free tool which can be used to visualize spatial data on satellite imagery, draw shapes (segments), and produce basic maps. Field Papers (http://fieldpapers.org/) is a non-technical free tool for creating paper field maps with an OpenStreetMap base layer. Planning Teams which do not have the support of a GIS team might consider using these two tools together to replicate most of the GIS team process.</p>
<p>Spreadsheets: MS Excel</p> 	<p>Excel is recommended for tracking changes and additions to PSU attributes throughout the process. Excel data can be easily joined to shapefiles. Tracking this information in Excel ensures that the Planning Team members (who do not use ArcGIS) have PSU attribute information readily available.</p>	<p>OpenOffice Calc (www.openoffice.org) is a good, free alternative to MS Office products.</p>


<p>MBtiles maker: Tile Mill</p>	<p>Tile Mill can be downloaded from https://tilemill-project.github.io/tilemill/ to make MBtile images for use in GeoODK. While Tile Mill provides OpenStreetMap data, the standard OpenStreetMap data is not updated regularly and thus may not reflect office mapping updates. We therefore provide instructions for using Tile Mill to create MBTiles from the ArcGIS OpenStreetMap baselayer, which reflects real-time updates from OpenStreetMap.</p>	
<p>GPX maker: My Geodata Converter</p>	<p>My Geodata Converter is an online tool available at: https://mygeodata.cloud/converter/shp-to-gpx to convert shapefiles into GPX format to be visualized in iD Editor and tablet apps.</p>	<p>QGIS has a built-in tool for converting shapefiles to GPX. ArcGIS has a user created toolbox extension for this task.</p>

Tablets

In the SUE surveys, each mapper-lister was assigned a Samsung Tab 3 or Tab 4 tablet. Only the relevant applications were enabled. The tablet was able to connect to WiFi, and enumerators-listers are provide data plans to allow for submission of forms daily. Table 7 summaries each app that was preloaded on tablets, and describes any alternative applications that were considered.

Table 8. Applications installed on tablets

Purpose: App	Description	Alternatives considered
Restriction: App Lock 	App lock enables only those applications which are relevant for enumeration and listing, and prevents enumerators-listers from changing key settings.	
Home screen: Nova Launcher 	Nova Launcher was used to customize the home screen of tablets.	
Submit daily reports: CamScanner 	CamScanner (www.camscanner.com) allows enumerator-listers to take a photo "scan" of paper documents at the end of each day and submit them to the Planning Team. Two scans should be sent each day: (1) field enumeration map (2) team tracking sheet	
Navigation: MAPS.ME 	MAPS.ME (http://maps.me) is used for navigation. While the app does not show satellite imagery, it does show major roads, bus stops, and landmarks for an entire country.	Here We Go (https://wego.here.com) is a good alternative to MAPS.ME for offline maps and navigation.
Mapping: OSMAnd + Paper Map 	OSMAnd (http://osmand.net/) was used as a reference while mapping buildings on paper map. The slide at the bottom of the screen can be used to change the transparency of the OpenStreetMap data to reveal satellite imagery below. If using the free version of the app, OSM data will only update every few weeks. If using the paid monthly subscription, OSM data will be updated continuously.	We considered use of Vespucci OSM Editor (http://vespucci.io/), an app that allows editing of buildings, streets, and points of interest offline in the field. We decided to not use this app because (a) it was difficult to see the screen in the field and to make edits on a small tablet screen, (b)

	<p>OSMAAnd allows the user to add points of interest to OSM, but does not allow editing of buildings or streets and thus was not used to collect geographic data. Please refer to "How to update paper map with OSMAAnd reference" for instructions.</p>	<p>did not allow a private option for recording shack/tent-residential buildings, and (c) it required about 3 times longer in the field compared to use of OSMAAnd+paper map.</p>
<p>Listing: GeoODK</p> 	<p>GeoODK (http://geoodk.com/) is form data collection tool used to collect listing and interview data. Each question appears on a separate page, and the user swipes left to move to the next question. We used the same listing form for one-stage and two-stage PSUs, but the form required fewer questions for one-stage dwelling listings.</p> <p>GeoODK displays the location of the tablet, and allows for the collection of a GPS coordinate at a specified location on an OSM map with each form. GeoODK either requires connection to a mobile phone network, or the use of MBtiles to work offline.</p> <p>The dwelling and household listing form was created in Excel and uploaded to the GeoODK server.</p>	<p>We also considered Open Map Kit and KoBo Collect for our listing and interview forms.</p> <p>Open Map Kit (http://openmapkit.org/) may be used as an alternative to GeoODK; rather than placing a GPS point over the listed building, the user selects the whole building feature from OSM and optionally labels it with a descriptive tags such as "concrete". This requires OSM building data to be accurate before field listing, however, which is why we decided against it.</p> <p>KoBo Collect (www.kobotoolbox.org) is not a good alternative GeoODK because it allows for collection of the GPS coordinate <i>where the tablet is located</i>, which is to sufficient to relocate buildings in dense urban settings.</p>

Review

Answer the questions below. To see if your answer is correct, go to Appendix D.

1. It is absolutely necessary to have at least one GIS specialist to conduct a gridded population survey?
 - a. TRUE. It is not possible to generate field maps without ArcGIS or QGIS.
 - b. FALSE. While a GIS specialist is recommended and very helpful in gridded population surveys, it is possible to generate field maps with Google Earth and Field Papers which are non-technical mapping tools.

2. Who should install and configure apps on the tablets for reporting, navigation, mapping, and listing?
 - a. Mapping-listing team
 - b. GIS team
 - c. Planning team

6. Logistics: Staff

Objectives:

By the end of this chapter, you will be able to:

- Hire mappers-listers and at least one GIS specialist
- Deliver the Mapper-Lister-GIS training

Outline:

- Hiring
- Training

Hiring

Mapper-lister. Mappers-listers should perform fieldwork in pairs. At least one person in each pair should have experience reading maps and using mobile apps. They might also have GIS experience or field data collection experience. These skill sets may be identified among undergraduate geography, GIS, or geoenvironment students. Please see “Enumerator-Lister Job Posting” in Appendix B for reference. It is also important that at least one person in each pair is comfortable conversing with members of the public and recording data. This skill set may be identified among public health professionals such as community health workers or professional data collectors.

In Nepal, we advertised for the mapper-lister position in a geo-engineering department, and accepted 14 applicants for the training. Using training exam scores, we will hire 12 mappers-listers, and 1 GIS specialist.

GIS specialist. The GIS specialist should have intermediate GIS skills developed from at least 1 year of classroom training and some applied GIS experience. The core skills include: generating maps, digitizing new features (points, lines, and polygons) and converting between spatial data formats (e.g. shapefile, GPX).

In Nepal and Bangladesh, The GIS Team was one employee embedded with the Planning Team. In Vietnam, the GIS Team was a hired consultant who worked with the planning team.

Training

Protocols. The Planning Team should closely review the Mapping-Listing-GIS Field Guide and accompanying powerpoint slides to prepare for the mapper-lister training. In Nepal, we offered the training over three days, but it required a fourth day of supplemental training for mappers-listers to gain practice approaching members of the public with the listing script. Feel free to improve and adopt to your local context.

Exam. Preparation of the practical exam materials may take several days, and may require help from the GIS Team. To create exam maps:

- Identify practice mapping-listing areas near the training facility
- In ArcGIS, create “answer” maps with OpenStreetMap base layer that include the exam area boundary, and markings (as line, box, or arrow shapes) indicating induced errors into OpenStreetMap
- Deliberately introduce these errors into OpenStreetMap using iD Editor
- Wait approximately 1 day for OpenStreetMap to update the base map and then create “exam” maps (with the induced errors) using ArcGIS.

Materials. The following materials should be printed for the training: 1 Mapping-Listing-GIS Field Guide and 1 set of powerpoint hand-outs per person. There are three listing activities that should be completed in pairs – be sure to print 1 copy of the assignment, and enough blank listing forms for everyone to practice listing.

For the Nepal Enumeration-Listing-GIS training for 14 people, we printed:

- 14 Enumeration-Listing-GIS Field Guides
- 14 sets of powerpoint handouts
- 7 back-up dwelling listing forms
- 7 back-up household listing forms
- 2 of each exam map (there were 7 exam sample areas)
- 1 of each exam map answers
- 1 dwelling listing easy exercise
- 1 dwelling listing complex exercise
- 1 household listing complex exercise

Review

Answer the questions below. To see if your answer is correct, go to Appendix D.

1. Mapping-listing pairs should have at least one person with experience reading maps and using mobile apps. Where might this type of employee be recruited?

- a. Undergraduate geography, GIS, or geo-engineering department
- b. Advanced GIS and engineering consulting firm
- c. General public

2. The mapper-lister training requires ____ days

- a. 1
- b. 3-5
- c. 10-12

3. The GIS required skills include:

- a. Generate maps, digitize new features, convert between file types
- b. Develop new mapping, listing, GIS protocols
- c. Export PSU locations to GPS units, and clean latitude/longitude coordinates

7. Logistics: Pre-field

Objectives:

By the end of this chapter, you will be able to:

- Form capable field teams, and assign them to PSUs
- Ensure fieldwork is feasible
- Prepare the materials needed for fieldwork

Outline:

- Set-up tools and infrastructure
- Determine final PSUs
- Assign mapping-listing teams
- Conduct pre-field segmentation

Set-up tools and infrastructure

Allow plenty of time (several weeks) to set up the GeoODK server, programme the GeoODK form, install and configure the apps on tablets, and to familiarize yourself with the software. The Planning Team might consider creating a few test sampling areas near your office and practice mapping with paper maps, and listing with the GeoODK form yourself to prepare for training and fieldwork.

Determine final PSUs

This activity happens after all primary sampling units (PSUs) are mapped in the office. In Nepal, 78 PSUs were generated and assigned the labels: “drop”, “main”, or “backup”. Of the 60 PSUs assigned to main, half were randomly assigned to the one-stage sample, and the other half to the two-stage sample, such that PSUs with non-permanent-looking building were evenly split between one-stage and two-stage samples.

Assignment of PSUs should be performed by at least one Planning Team member who is intimately familiar with the different areas of the survey coverage area, with support from a GIS team member.

Review PSUs with GIS Specialist

Step 1: Have the GIS specialist visualize each PSU in ArcGIS with OpenStreetMap (updated during office mapping) and satellite imagery as base layers. Decide if any PSUs will be dropped because no habitable buildings are visible inside of the PSU. Figure 16 shows a PSU that was dropped in Nepal because the only building inside the PSU or on the boundary was a Diesel generator plant, which is not expected to have any inhabitants.



Figure 16. Example dropped PSU

Set-up PSU assignment worksheet

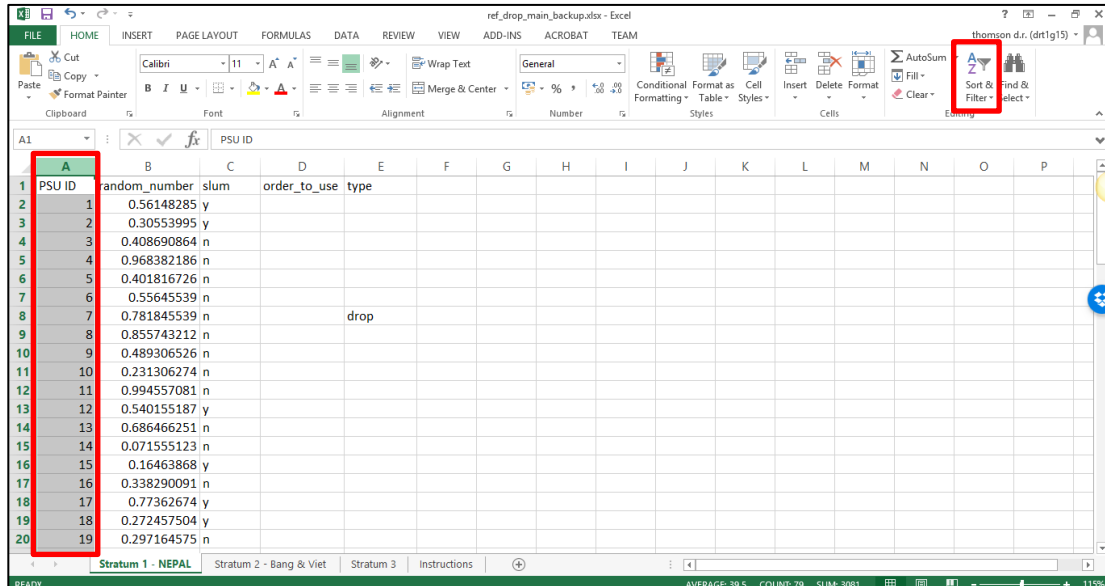
Step 2: Create assign_psus.xlsx document with a column for unique PSU ID, and a column that has a random number between 0 and 1 for each PSU. If you have used the random number function in Excel [=rand()], this random number will dynamically regenerate when you make changes to the document.

Assign “drop” PSUs

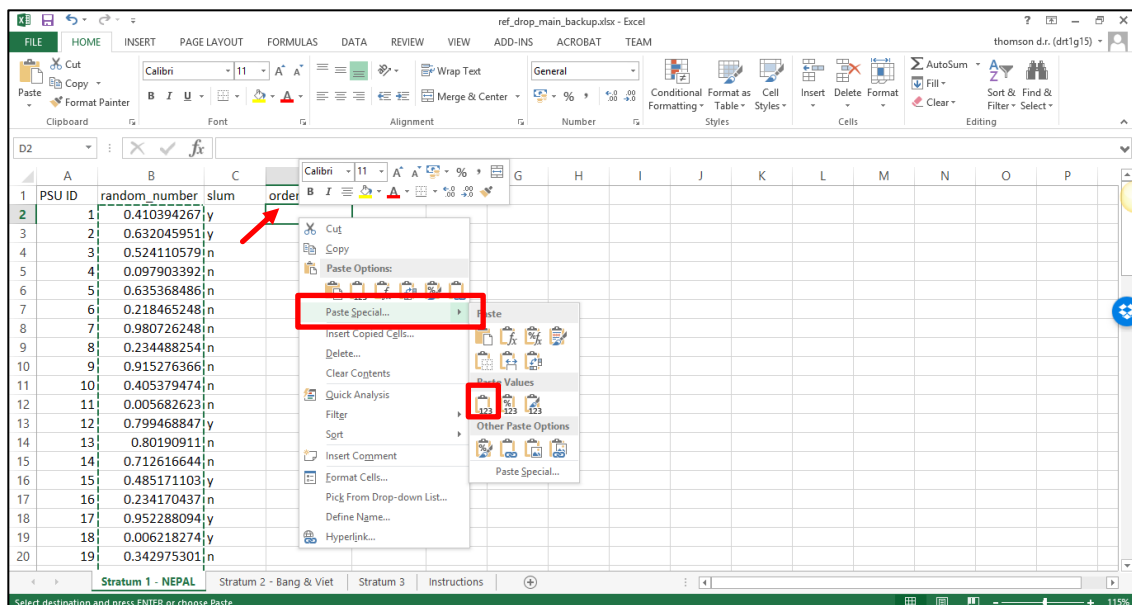
Step 3: If any PSUs need to be dropped, type "drop" next to the PSU ID in the "type" column. We will ignore these PSUs moving forward as if they were never selected.

Assign "main" and "backup" PSUs

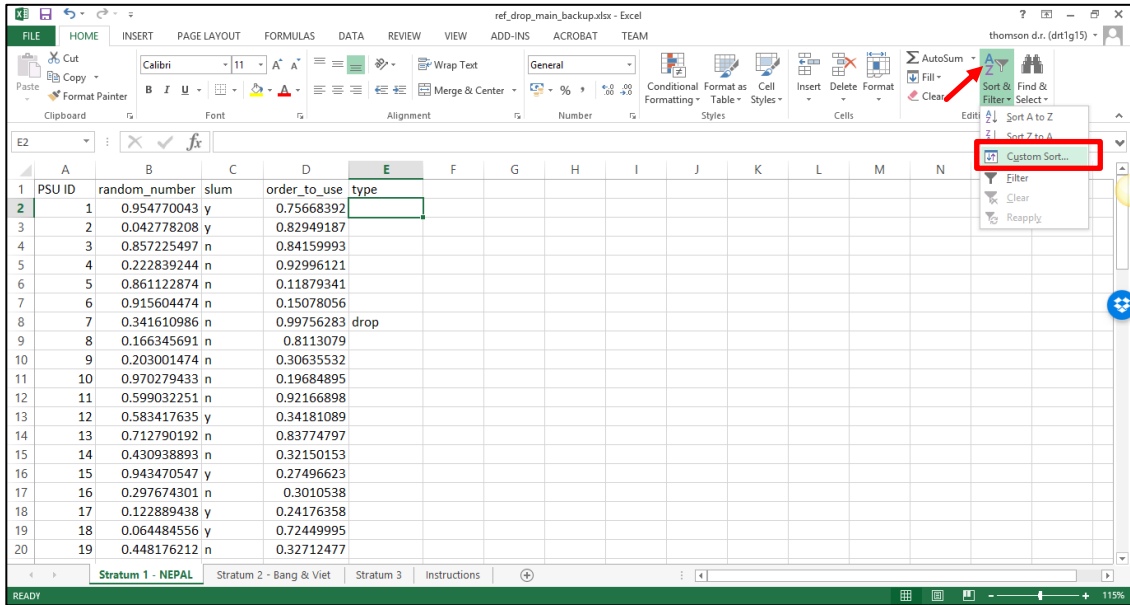
Step 4: Sort PSU IDs from smallest-to-largest (ascending)



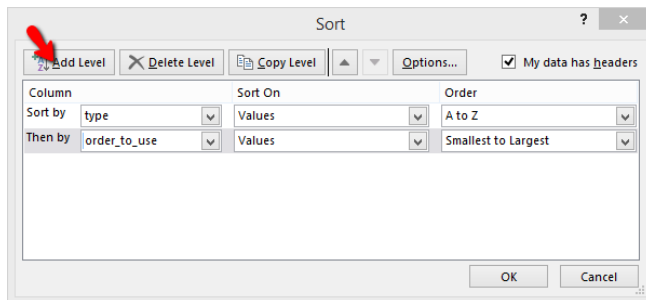
Step 5: To create a permanent random number for each PSU, copy the generated numbers, right-click the first cell in the "order_to_use" column, and "Paste Special" the values.



Step 6: Open the custom sort tool



Step 7: Sort "type" from A to Z. Use the "Add Level" button to add "order_to_use", and sort it from smallest-to-largest (ascending)



Step 8: Assign "main" to the first 60 PSUs (or 20 in each Vietnam and Bangladesh), then assign "backup" to the remaining PSUs.

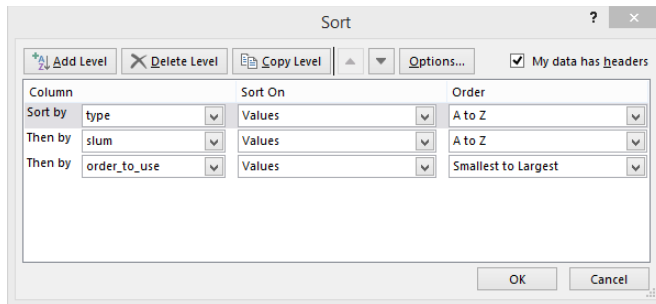
	A	B	C	D	E
1	PSU ID	random_number	slum	order_to_use	type
2	7	0.446662176	n	0.99756283	drop
3	30	0.889940122	y	0.00717277	main
4	57	0.214077194	n	0.01796553	main
5	47	0.61629132	n	0.03318193	main
6	67	0.588348173	y	0.03658106	main
7	27	0.373119568	y	0.06482646	main
8	60	0.329924879	n	0.07121702	main
9	66	0.157775219	n	0.07403985	main
10	21	0.753285555	n	0.10481599	main
11	40	0.838168125	n	0.11680935	main

.....

69	3	0.321148066	n	0.84159993	backup
70	28	0.975441573	n	0.84522933	backup
71	69	0.849632768	n	0.85518371	backup
72	51	0.786982008	y	0.86659421	backup
73	25	0.063400778	n	0.89044594	backup
74	11	0.383474923	n	0.92166898	backup
75	4	0.108745081	n	0.92996121	backup
76	41	0.795126438	n	0.93498622	backup
77	35	0.252534852	n	0.93650043	backup
78	58	0.291262881	y	0.9428845	backup
79	75	0.726674013	n	0.98665139	backup

Assign “one-stage” and “two-stage” PSUs (Nepal only)

Step 9: Open the custom sort tool, and sort “type” from A to Z. Use the “Add Level” button to add “slum”, and sort from A to Z. Finally use add “order_to_use”, and sort it from smallest-to-largest (ascending). The sort order matters.



Step 10: In this step we will reassign half of the “main” PSUs with slum areas (group of informal-looking buildings) to the one-stage sample, and reassign the other half of slum PSUs to the two-stage sample. Scroll to the bottom of your sorted PSU list. Count the number of PSUs with slum areas, and divide by 2 (if it is an odd number, adding the extra slum PSU to the one-stage sample). The use of colour to highlight cells might be useful. Reassign the one-stage PSUs first, then two-stage PSUs second.

	A	B	C	D	E	F	G
61	32	0.787171946	n	0.72322921	main		
62	62	0.906476669	n	0.74389123	main		
63	64	0.937229523	n	0.74898747	main		
64	43	0.989948884	n	0.7570458	main		
65	30	0.281499669	y	0.00717277	one		
66	67	0.991658301	y	0.03658106	one	15 / 2 = 7.5	
67	27	0.578703072	y	0.06482646	one	one-stage = 8	
68	17	0.34863834	y	0.24176358	one	two-stage = 7	
69	15	0.972423828	y	0.27496623	one		
70	12	0.100120875	y	0.34181089	one		
71	22	0.958783333	y	0.38273915	one		
72	56	0.46925681	y	0.51163964	one		
73	38	0.054781428	y	0.58972965	two		

Step 11: Now scroll to the non-slum “main” PSUs, and do the same – reassign half to the one-stage sample and half to the two-stage sample. If there is an odd number, add the extra non-slum PSU to the two-stage sample. Reassign the one-stage PSUs first, then two-stage PSUs second.

20	57	0.717177103	n	0.01796553	one	
21	47	0.080457439	n	0.03318193	one	45 / 2 = 22.5
22	60	0.914160825	n	0.07121702	one	one-stage = 22 PSUs
23	66	0.26285729	n	0.07403985	one	two-stage = 23 PSUs
24	21	0.92194996	n	0.10481599	one	
25	40	0.324148546	n	0.11680935	one	
26	5	0.691387838	n	0.11879341	one	
27	36	0.188863164	n	0.1325218	one	
28	55	0.2197212	n	0.14633356	one	

Check your work

This is an important moment in the sample process. Let us summarize our sample, and ensure we didn't make a mistake.

Step 12: Go to the Insert tab. Select the entire data table and create a pivot table. Select “OK” to the defaults in the pop-up window.

The screenshot shows the Excel interface with the 'INSERT' tab active. The 'PivotTable' icon in the 'Tables' group is highlighted with a red box. The 'Create PivotTable' dialog box is open, displaying the following options:

- Choose the data that you want to analyze:**
 - Select a table or range
 - Table/Range: 'Stratum 1 - NEPAL!\$A:\$E'
 - Use an external data source
- Choose where you want the PivotTable report to be placed:**
 - New Worksheet
 - Existing Worksheet
 - Location: [Empty text box]
- Choose whether you want to analyze multiple tables:**
 - Add this data to the Data Model

The background shows a data table with columns A through E:

	A	B	C	D	E
1	PSU ID	random_number	slum	order_to_use	type
2	70	0.500822892	n	0.77066664	backup
3	34	0.131109115	n	0.77300779	backup
4	49	0.767779384	n	0.80326422	backup
5	8	0.891711206	n	0.8113079	backup
6	13	0.084372013	n	0.83774797	backup
7	3	0.362383837	n	0.84159993	backup
8	28	0.94607244	n	0.84522933	backup
9	69	0.048579368	n	0.85518371	backup
10	25	0.103152648	n	0.89044594	backup

Step 13: Drag “type” to the ROWS area, and “PSU ID” to the VALUES area. This will produce a count of PSUs by type.

Check that (a) all PSUs are assigned to a type, and (b) that there is the correct number of PSUs in each type.

If you identify any errors, return carefully to each of the above steps. If the summary table is correct, then you are done! You know which PSU IDs will be in the “main” sample (Bangladesh/Vietnam), or in the “one-stage” and “two-stage” samples (Nepal).

Row Labels	Count of PSU ID
backup	17
drop	1
one	30
two	30
(blank)	
Grand Total	78

PivotTable Fields

Choose fields to add to report:

- PSU ID
- random_number
- sl...
- order_to_use
- type

Drag fields between areas below:

ROWS: type

VALUES: Count of PSU ...

Defer Layout Update UPDATE

Team formation

Mappers-listers work in the same pair throughout their contract, so the Planning Team should carefully consider how teams are formed, and where they are assigned to work.

In Nepal, we laid out the photo ID cards for all hired mappers-listers, and arranged them into pairs using the following considerations:

- Identified the strongest individuals to act as team leads
- Identified the weakest individuals to be paired with strong partners
- Interpersonal relationships were considered when pairing enumerator-listers to try to avoid conflict.
- Home locations of mappers-listers were also considered. When possible, we paired people who lived nearer to each other, and tried to assign them to PSUs accessible to their home region of the city.
- In Kathmandu Valley, there is one region dominated by the Newari ethnic group, and thus we considered a team comprised of two Newari speakers to work in this area.

Separate from team formation, someone on the Planning Team sat with a GIS specialist to group all of the final PSUs by region of the city.

At the end of the office enumeration period, the planning team held a meeting with all mappers-listers to (a) give them their team assignments, and (b) discuss as a large

group which team would work in each region of the city. While the Planning Team already had ideas about team regional assignments, this was discussed as a group before being finalized to ensure transparency, and to allow mappers-listers to raise any additional considerations that could improve logistics of the fieldwork.

Pre-field segmentation

Segmentation is an important step to ensure feasible field work. Someone from the Planning Team should review all sampling areas after office mapping and before field mapping to identify PSUs that appear to have more than 30 (one-stage) or 300 (two-stage) dwellings.

We recommend that the Planning Team member sits with a GIS specialist during this process to discuss where to manually divide the PSUs such that all segments have approximately equal-population, using streets, rivers, and other natural features to draw segment boundaries. Remember that segments should not divide buildings or properties. The number of segments made in each sampling area should be tracked, incorporated into the sample weights, and the sampling weights should be used in the survey data analysis. Refer to the Mapper-Lister-GIS manual for steps to track number of segments created per PSU, and random selection of one segment per PSU.

Also remember that one-stage sampling areas may be segmented a second time after field mapping based on number of dwellings listed, so in this step, air on the side of caution (more buildings) when segmenting in commercial areas in case buildings have few dwellings.

Review

Answer the questions below. To see if your answer is correct, go to Appendix D.

1. In the SUE surveys, PSUs are assigned which of the following labels:

- a. Drop, main, backup
- b. A, B, C, D
- c. Pre-field, fieldwork, post-field

2. In the SUE surveys, mapping-listing teams were comprised of pairs and travelled to the field directly from home on most days. What considerations did the Planning Team give to team formation?

- a. Individual mapping-listing skill level
- b. Inter-personal relationships observed during the training
- c. Home location
- d. Spoken language skills
- e. All of the above

3. The Planning Team should help conduct pre-field segmentation when:

- a. One-stage PSUs appear to exceed 30 dwellings in satellite imagery
- b. Two-stage PSUs appear to exceed 300 dwellings in satellite imagery
- c. A and B

8. Logistics: Mapping-Listing Fieldwork

Objectives:

By the end of this chapter, you will be able to:

- Develop and use reports to identify errors and poor quality data in real-time
- Support mappers-listers to identify and overcome field challenges

Outline:

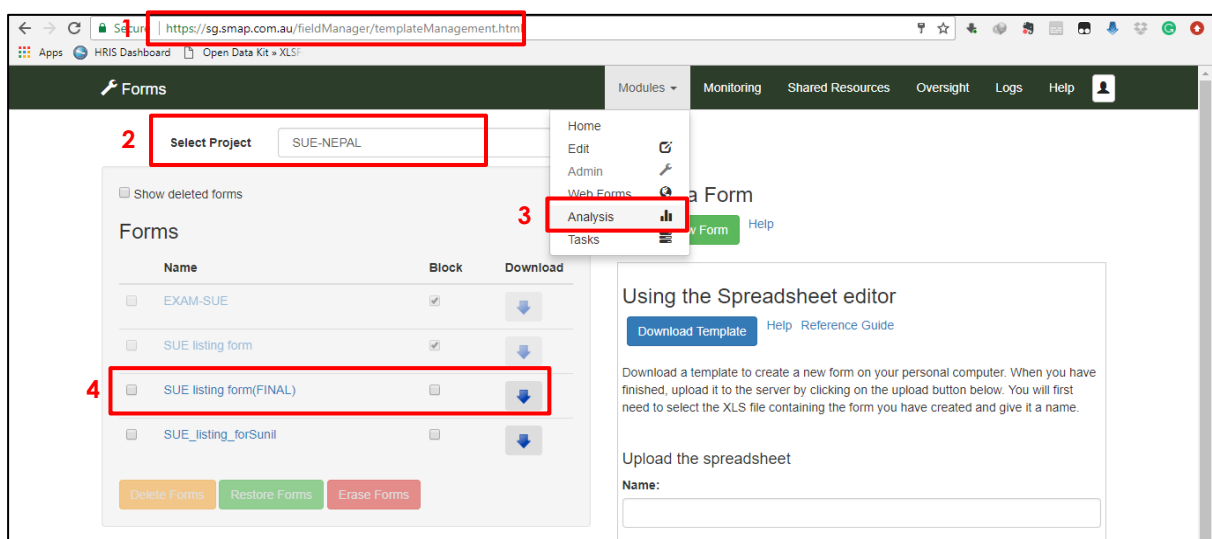
- Quality control of mapping-listing data
- Office meetings
- Field visits

Quality Control

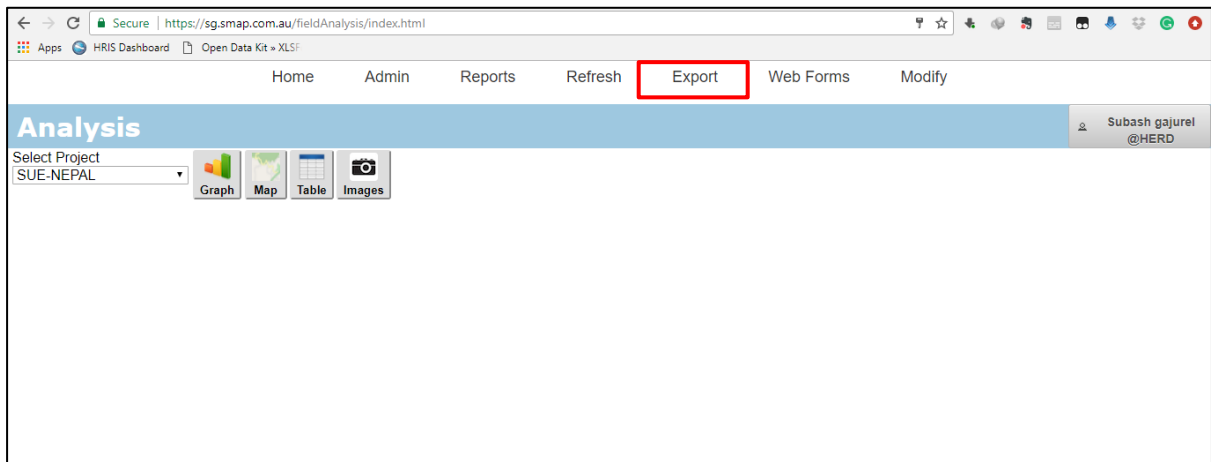
Once fieldwork begins, the Planning Team will begin quality checking data and process in the office and in the field. We recommend four daily quality checks on the enumeration maps and listing data. To perform these checks, export the listing data to excel and to shapefile. These steps may be different depending on your server set-up, though the final data quality checks should be similar.

Export listing data from GeoODK to Excel

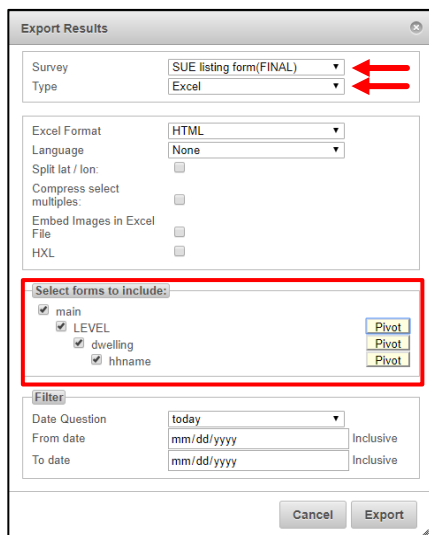
Steps 1-4: Log into sg.smap.com.au, and select the relevant SUE project. Under the “Modules” tab in the menu bar, select “Analysis”, and then select the form for which you would like to export data.



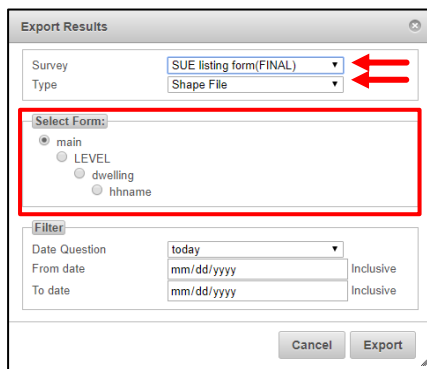
Step 5: The Analysis screen will open. Select “Export” from the menu bar.



Step 6: To export the raw data to Excel, select the survey form to be exported, choose type = “Excel” and ensure that all levels of the form are selected with “pivot” format.



Step 7: To export the raw data to a shapefile, select the survey form to be exported, choose type = "Shapefile" and select the "main" level of the survey (at this level, one record = one building).



Step 8: Use random sample method in stage 2 to select 20 HH in one PSU. Create one column named as Sampled HH. Selected HH are then denoted from 1 to 20.

Example:

Sampling (Rand No)	Sampling (Order No)	Sampled HH	New ID	N.SN.	SN	prikey	enu	enu1	enu_name	psu1	PSU	STAGE	ADDRESS	BUILDING TYPE	LEVEL_NO	DWELLING _NO	HOUSEHOL D_NO	HC D_NO
0.3363946	0.0032987	1	96	1472	7169	1277	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.8531327	0.0058335	2	82	1458	7129	1253	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.4611582	0.0228625	3	99	1475	7186	1285	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.9735671	0.0454202	4	14	1096	5179	924	2	2	Bhandari	2	PSU = 2	STAGE = 2	Chunikhel -	TYPE =	1	NO = 1	D NO = 1	D1
0.1790970	0.0484348	5	6	1072	5074	913	2	2	Bhandari	2	PSU = 2	STAGE = 2	Chunikhel -	TYPE =	1	NO = 1	D NO = 1	D1
0.1077857	0.0530726	6	98	1474	7184	1284	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.9717970	0.0668172	7	34	1298	6488	1173	2	2	Bhandari	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.5965083	0.0904870	8	113	1944	9264	1546	2	2	Bhandari	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.0779803	0.0914584	9	36	1300	6492	1175	2	2	Bhandari	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.3919478	0.0915395	10	92	1468	7155	1269	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.1792314	0.0924415	11	110	1941	9260	1543	2	2	Bhandari	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.6366150	0.1001691	12	119	1952	9285	1556	2	2	Bhandari	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.3951800	0.1119299	13	40	1304	6500	1178	2	2	Bhandari	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.6838064	0.1202061	14	56	1320	6535	1197	2	2	Bhandari	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.8607147	0.1224878	15	94	1470	7165	1275	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.1897699	0.1312518	16	49	1313	6518	1187	2	2	Bhandari	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.4448609	0.1360131	17	93	1469	7163	1274	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.3104217	0.1463339	18	105	1936	9248	1535	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.0955099	0.1874366	19	9	1078	5092	918	2	2	Bhandari	2	PSU = 2	STAGE = 2	Chunikhel -	TYPE =	1	NO = 1	D NO = 1	D1
0.2258310	0.1879317	20	102	1478	7194	1290	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.9681467	0.1975004	21	32	1296	6483	1171	2	2	Bhandari	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.6302891	0.2032355	22	74	1450	7113	1239	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1
0.6986942	0.2050148	23	80	1456	7122	1247	3	3	Aryal	2	PSU = 2	STAGE = 2	Baluwa -3	TYPE =	1	NO = 1	D NO = 1	D1

Step 9: Filter excel data by stage = 1 and Building Type = C and q2 = 0. Deleted all rows. One column named as sampled HH is created. All remaining dwelling are then denoted from 1 to N

Where C = Commercial

q2 = Total number of dwellings on level X

N = Number of rows

Step 10: In the Excel file, create a new field called unique_dwelling by concatenating PSU ID, Sampled HH, Building ID, Level ID, Dwelling ID and HH No for stage 2 and for stage 1 by concatenating PSU ID, Sampled HH, Building ID, Level ID, Dwelling ID . For example:

=CONCATENATE (P12, ".", A12, ".", AC12, ".", AE12)

Pivot Tables

Pivot tables in Excel are used to create reports. If the Planning Team is not familiar with pivot tables, we strongly recommend watching YouTube Videos and learning to use pivot tables, rather than summarizing raw GeoODK exports manually. The raw listing data in Excel has one record per dwelling (one-stage PSUs) or household (two-stage PSUs). Please see figure 3 for an example of the raw listing data from GeoODK. Note that some fields are hidden for simplicity, and other fields are anonymized with ***** to protect the identity of SUE-Nepal respondents.

Data dictionary:

- prikey: unique building ID generated by GeoODK
- enu: enumerator-lister ID
- enu1: confirmation of enumerator-lister ID
- enu_name: enumerator-lister's name
- psu1: PSU ID
- psu2: confirmation of PSU ID
- stage: whether PSU is one-stage or two-stage sample (Nepal only)
- the_geom: latitude-longitude coordinate of building location
- building: building type where R=residential, C=commercial/institutional, M=mixed, V=vacant/construction/demolished, and S=shack/tent
- q1: total number of levels
- area: area name
- ward: ward number
- tole: tole name
- info: source of information where R=resident, N=neighbour, and L=lister
- image: link to photo of the building or building gate
- detail: detailed description of the building
- building_level: building level ID where 1=ground level
- q2: total number of dwellings on level X
- dwelling_no: dwelling ID on level X
- hhn: total number of households in dwelling X on level X
- hh_no: household ID in dwelling X on level X
- q3: head of household name (use "999" if household name could not be obtained, and provide a detailed explanation about why name was not obtained in description field).
- unique_dwelling: a unique dwelling ID comprised of PSU, Building, Level, and Dwelling IDs

	A	M	N	O	P	Q	R	S	T	U	V	W	X	Y	AA	AB	AC	AD	AE	AF	AG	AH	AI
1	prikey	enu	enu1	enu_name	psu1	psu2	stage	the_geom	building	q1	area	ward	tole	info	image	detail	building_level	q2	dwelling_no	hhn	hh_no	q3	unique_dwel
2	3	4	4	Bijay Dhakal	19	19	2	*****	M	4	Bishal nagar chok	5	****	R	*****	Rising photo studio, dhMla bary and bakery shop,eye clinic, house at junction	1	0			1	****	19.3.1.
3	3	4	4	Bijay Dhakal	19	19	2	*****	M	4	Bishal nagar chok	5	****	R	*****	Rising photo studio, dhMla bary and bakery shop,eye clinic, house at junction	2	1	1	1	1	****	19.3.2.1
4	3	4	4	Bijay Dhakal	19	19	2	*****	M	4	Bishal nagar chok	5	****	R	*****	Rising photo studio, dhMla bary and bakery shop,eye clinic, house at junction	3	0			1	****	19.3.3.
5	3	4	4	Bijay Dhakal	19	19	2	*****	M	4	Bishal nagar chok	5	****	R	*****	Rising photo studio, dhMla bary and bakery shop,eye clinic, house at junction	4	0			1	****	19.3.4.
6	4	4	4	Bijay Dhakal	19	19	2	*****	M	5	Bishal nagar chok	5	****	R	*****	Nischal traders,shree bm shyamk traders,narendra trade link,pharmacy,purple color shutter	1	0			1	****	19.4.1.
7	4	4	4	Bijay Dhakal	19	19	2	*****	M	5	Bishal nagar chok	5	****	R	*****	Nischal traders,shree bm shyamk traders,narendra trade link,pharmacy,purple color shutter	2	1	1	1	1	****	19.4.2.1
8	4	4	4	Bijay Dhakal	19	19	2	*****	M	5	Bishal nagar chok	5	****	R	*****	Nischal traders,shree bm shyamk traders,narendra trade link,pharmacy,purple color shutter	3	0			1	****	19.4.3.
9	4	4	4	Bijay Dhakal	19	19	2	*****	M	5	Bishal nagar chok	5	****	R	*****	Nischal traders,shree bm shyamk traders,narendra trade link,pharmacy,purple color shutter	4	0			1	****	19.4.4.

Figure 17. Example raw listing data export from GeoODK to Excel

Check 1: Completeness of listing entries. Review the raw listing data, and spot-check the following for each field team:

- Description field (detail): Includes a street and house number, if available. Includes distinct features, such as colour or material of walls, or location and number of water tanks and solar panels.
- Head of household name (q3): Full names are entered, and spelled correctly.

Check 2: Dwelling/household count and type report. The dwelling/household report ensures that teams are working in the PSUs they say they are working in, and they are entering PSU IDs correctly. The summary report also allows you to monitor the speed at which each team is working. An example summary report is shown in Figure 18, and the pivot table set-up for this report is shown in figure 19.

	A	B	C	D	E	F	G
1	dwelling_count (sums)						
2							
3	Count of column L						
4	Row Labels	C	M	R	S	V	Grand Total
5	1	19	176	134	2		331
6	1		3	9			12
7	4		4	2			6
8	13		26	11			37
9	16	2	29	16			47
10	20	1	27	4			32
11	23		9	26	2		37
12	35	11	15	6			32
13	40			9			9
14	41	1	24	1			26
15	49	3	19				22
16	55	1	3	23			27
17	62		17	25			42
18	123			2			2
19	2	33	600	1427	32	7	2099
20	2		17	89	19	3	128
21	6	1	11	154			166
22	9	11	136	326	4	2	479
23	14	5	1	88			94

Figure 18. Example dwelling/household count and type report

Drag fields between areas below:

<p>FILTERS</p> <p>dwelling_no</p>	<p>COLUMNS</p> <p>building</p>
<p>ROWS</p> <p>stage</p> <p>psu1</p>	<p>VALUES</p> <p>Count of dwelling_no</p>

Figure 19. Pivot table configuration for dwelling/household count and type report

Counting the dwelling ID field (dwelling_no) in the raw data results in a count of dwellings (one-stage PSUs) and a count of households (two-stage PSUs). In Nepal, GeoODK duplicated records in the database, requiring additional data cleaning. We found that dwelling ID field (dwelling_no) was blank for all duplicated records. Thus we were able to exploit this to filter out duplicate records where dwelling_no = blank.

In the dwelling/household configuration report, look for the following:

- PSU IDs are correct: In Figure 18, the one-stage PSU "123" did not exist in our survey, and the two-stage PSU "9" was not due to be visited by a field team. These are both data entry errors. The Nepal Planning Team was able to spot both errors, and immediately called the respective field teams to make corrections in their future listing forms. Errors in the form data were manually cleaned.
- Monitor total number of dwellings/households per PSU: In Figure 18, most one-stage PSUs have 20+ dwellings, as expected. PSU 4 and 40, however, had only 6 and 9 dwellings, respectively. The Planning Team checked the field maps for these PSUs and confirmed that one PSU was mostly occupied by a temple, and the other was mostly occupied by a water treatment plant. Thus, the low dwelling counts made sense. The planning team still called the field teams in both PSUs to confirm description of the PSUs from the field.
- Monitor the speed at which listing is occurring. In Figure 4, the two-stage PSU "6" has 166 households listed and it took several days. The Planning Team monitored work on this PSU in real-time; after two full days, the team had only listed 90 households, which was many fewer households than other teams had listed in two-stage PSUs. The Planning Team called the field team to see what difficulties they might be facing in terms of commute times, refusals, their process, etcetera.
- Building-residence scenarios. It is essential that enumerators are holding detailed conversations with residents and neighbours to list all households and dwellings accurately. Generally, the population will live in residential and mixed buildings, but we expected a certain number of guards, cleaning staff, and other workers to stay long-term in commercial/institutional buildings, shacks/tents, and even vacant/construction/demolished buildings. From this report, the Planning Team was able to identify PSUs and teams which list only traditional households, in order to prompt follow-up conversations with the teams and flag potential field supervision visits.

Check 3: Households per dwelling report (two-stage only). The report of households per dwelling helps to flag teams that are not holding detailed conversations with residents and neighbours to understand living arrangements. It is very common in Kathmandu and other Asian cities for multiple households to share a dwelling, thus we expected to find some dwellings with multiple households in each two-stage PSU.

In figure 20, the example households per dwelling report for PSU “6” shows that none of the dwellings in this PSU had multiple households listed. The Planning Team saw this, and called the field team to ask what the PSU was like and why multi-households dwellings were not identified. The field team reported that this PSU was located in a very wealth community where most buildings were large, gated single family homes. The Planning Team asked whether these residences had guards and cleaning staff, and the planning team confirmed yes, and that they were listing guards and cleaning staff in separate dwellings from the main residence. This was the correct protocol.

	A	B
1	stage	2
2	dwelling_no	(Multiple Items)
3		
4	Row Labels	Count of hh_no
5	2	119
6	6	164
7	6.10.1.1	1
8	6.131.1.1	1
9	6.131.2.1	1
10	6.131.3.1	1
11	6.134.1.1	1
12	6.135.1.1	1
13	6.136.1.1	1
14	6.136.3.1	1
15	6.136.4.1	1
16	6.137.1.1	1
17	6.138.1.1	1

Figure 20. Example households per dwelling report (two-stage only)

Drag fields between areas below:

FILTERS	COLUMNS
stage	
dwelling_no	
ROWS	VALUES
psu1	Count of hh_no
unique_dwll	

Figure 21. Pivot table configuration of households per dwelling report

Check 4: Comparison of listing data and paper field map. The location of buildings, building type, and number of building levels should be consistent in the listing data and paper field map.

Figure 22 shows each building latitude/longitude coordinate in example PSU mapped in ArcGIS, with building type and building levels marked as a label. Figure 23 shows the paper field maps drawn for the same example PSU.



Figure 22. Building latitude/longitude coordinates mapped for example PSU in ArcGIS

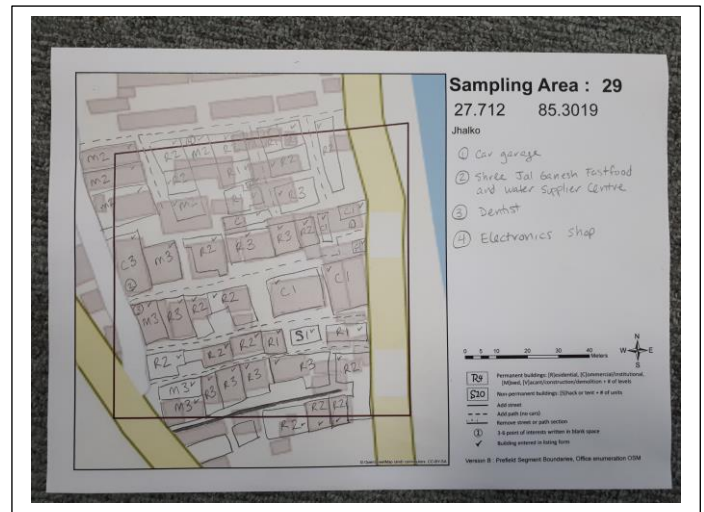


Figure 23. Paper field map for Mapping-Listing phase in example PSU

The GIS Specialist should map the shapefile export from GeoODK and:

- Visually compare the map generated from listing output against the photo scan of the field paper map to ensure that building type and number of levels are consistent. The Planning Team should call the field team and discuss if any discrepancies are identified.
- Check that listing latitude/longitude coordinates are located on top of building features. It is very important that interviewers are able to return to the correct building to perform interviews.

Office meetings

For the first few weeks of fieldwork, mappers and listers attended meetings at the office to update OpenStreetMap, and to debrief about their field experience. In these meetings, the Planning Team addressed questions, raise quality control issues observed in submitted daily maps, listing uploads, or team tracking sheets.

In Nepal, the first office meeting was extremely valuable to share experience across teams, and identify small processes that teams had developed themselves to streamline fieldwork. For example, one team was marking buildings in the paper field map with an asterisks if no one was available to identify buildings to return to and list later. After discussing, all teams decided to adopt this process.

Requiring enumeration-listing teams to come to the office on the same day at the same time is less important after the first meeting. In Kathmandu, it often took teams 1-2 hours to get to sampling areas from home, and an additional 1-2 hours to travel to the office. Thus, the Planning Team let each team finish work in two PSUs between office visits, even if it meant team were out of sync on office visit days.

Field visits

During the first two weeks of fieldwork, the Planning Team visited each mapping-listing team in the field at least once each. During field visits, the mapping-listing team was asked to show the paper field map, and to list one household whilst being supervised. Feedback on the following was provided by the Planning Team:

- Presentation – Is the team wearing badges, and acting professional?
- Interactions – How is the team approaching residents and neighbours? How are they fielding questions, rejection, suspicion, and curiosity? Provide guidance that can improve the quality of these interactions.
- Thoroughness – How thoroughly is the team identifying buildings? Are they recording building type for all buildings? Are they including all paths? Are they probing residents and neighbours to fully understand living arrangements? Are they recording the full name of head of household in the listing form?
- Accuracy – Review maps to ensure building locations and approximate shapes are accurate. Observe a listing and ensure the team is interpreting “dwelling” and “household” definitions correctly, and entering the correct number of levels, dwellings, and households in GeoODK.

Here are materials that are helpful for the Planning Team to bring on field visits:

- ✓ Cell phone
- ✓ Contact list for all enumerator-listers
- ✓ Rain coat/umbrella/sunscreen
- ✓ Water
- ✓ ID badge
- ✓ Tablet with all PSU boundaries loaded
- ✓ Paper maps of sampling area to be visited
- ✓ Exam evaluation forms (a good reference for checking quality)
- ✓ Informational handouts (x5) when you repeat a random listing
- ✓ Dwelling/household count and type report (created from listing data)
- ✓ Households per dwelling report (two-stage only) (created from listing data)

Review

Answer the questions below. To see if your answer is correct, go to Appendix D.

1. Quality control of data and processes is done by the Planning Team via:
 - a. Daily monitoring of listing data and maps submitted by mapping-listing teams
 - b. Routine face-to-face office meetings with mappers-listers
 - c. At least one field visit by the Planning Team to each mapping-listing team in the early days of fieldwork
 - d. All of the above

2. Recommended quality control reports check for:
 - a. Completeness of listing entries
 - b. Count of dwellings/households and their building type (e.g. mixed)
 - c. A and B

3. The GIS Specialist helps to compare paper field maps and building latitude/longitude coordinates collected in the listing. The GIS Specialist is looking for:
 - a. Consistency between the paper field map and lat/long locations
 - b. Lat/long coordinates are located on top of buildings when visualized on top of satellite imagery or OpenStreetMap
 - c. A and B

9. Logistics: Post-field

Objectives:

By the end of this chapter, you will be able to:

- Select the final samples
- Generate final maps needed for interviews
- Calculate sample design weights

Outline:

- Select household sample (two-stage only)
- Post-field segmentation (one-stage only)
- Prepare map atlases for interviews
- Prepare household listing for interviews
- Sample weights

Select household sample (two-stage only)

After field mapping and listing, the Planning Team selects households within two-stage sampling areas to be interviewed. According to the sample size calculation, 20 households should be selected per cluster. Households were selected systematically using a sampling interval.

Step 1: In Excel, list all two-stage households ordered from 1...n. Add the following columns, calculated for each PSU: PSU ID, total number of listed households in PSU, sampling interval in PSU, and random number for PSU.

Sampling interval in PSU = total number of listed households in PSU ÷ 20,
rounded to the nearest whole number

Random number is between 1 and 20 (inclusive)

Step 2: Add a column that identifies whether the household will be sampled or not. Start at household 1 in the PSU. Identify the household that corresponds to the random number between 1 and 20. Mark this household to be included in the sample. Now count down the list, and mark for inclusion every household that falls on the PSU sample interval.

In the below example (Figure 24), PSU 1 has 179 households. The sampling interval is $179 \div 20 = 8.95$ which rounds to 9. The random number 4 was determined using the excel function: " =RANDBETWEEN(1,20)". Selected households are marked in the "sample" column with 1, and highlighted.

household_id	psu_no	bdlg_no	level_no	dwell_no	hh_no	hh_per_psu	interval	rand_no	sample
001.01.1.01.01	1	1	1	1	1	179	9	4	0
001.01.1.01.02	1	1	1	1	2	179	9	4	0
001.01.2.01.01	1	1	2	1	1	179	9	4	0
001.01.2.01.02	1	1	2	1	2	179	9	4	1
001.02.1.01.01	1	2	1	1	1	179	9	4	0
001.03.1.01.01	1	3	1	1	1	179	9	4	0
001.04.1.01.01	1	4	1	1	1	179	9	4	0
001.04.1.01.02	1	4	1	1	2	179	9	4	0
001.04.1.02.01	1	4	1	2	1	179	9	4	0
001.05.1.01.01	1	5	1	1	1	179	9	4	0
001.06.1.01.01	1	6	1	1	1	179	9	4	0
001.07.1.01.01	1	7	1	1	1	179	9	4	0
001.07.2.01.01	1	7	2	1	1	179	9	4	1
001.07.3.01.01	1	7	3	1	1	179	9	4	0
001.08.1.01.01	1	8	1	1	1	179	9	4	0
001.08.1.01.01	1	8	1	1	1	179	9	4	0
001.08.1.01.01	1	8	1	1	1	179	9	4	0
001.09.1.01.01	1	9	1	1	1	179	9	4	0

Figure 24. Example selection of households sample in two-stage PSUs

Post-field segmentation (one-stage only)

Post-field segmentation will be needed in one-stage PSUs to ensure that interview teams find approximately 20 households in each PSU. PSUs with more than 25 dwellings should be segmented. Decisions about where to draw post-field segments should be made by a Planning Team member and implemented by the GIS specialist.

To show dwelling counts per building in ArcGIS and decide on post-field segmentation, the Planning Team must provide the GIS Specialist with a shapefile export from the listing database. Follow instructions in “Logistics: Fieldwork” chapter, in the section called “Export listing data from GeoODK”.

As with pre-field segmentation, someone from the Planning Team should review all one-stage sampling areas with the GIS specialist and discuss where to manually segment PSUs such that segments follow natural features such as roads and rivers and have equal number of dwellings.

Each one-stage PSU should have 15-20 dwellings. So if a PSU has 47 total dwellings listed, then 3 equal size segments of 15 or 16 dwellings each should be created ($47 \text{ dwellings} / 3 \text{ segments} = 15.7 \text{ dwellings per segment}$). Refer to the Enumerator-Lister-GIS manual for steps to track number of segments create per PSU, and random selection of one segment to represent the PSU.

Download permanent record of OpenStreetMap data

After field enumeration and updating OpenStreetMap, the GIS team downloads a snap-shot of OSM data to use in the next phase of the survey – interviews. This is a VERY important step, as OpenStreetMap is a dynamic public dataset, and it is possible that others could change the building and street data that your team worked hard to update in sampling areas. Instructions are provided in the Mapper-Lister-GIS manual.

Map shack/tent-residential buildings offline

The downloaded permanent record from OpenStreetMap should be updated in ArcGIS or QGIS with any shack/tent-residential structures identified in PSUs. These should NOT have been entered in OpenStreetMap previously. Adding and modifying polygon features in ArcGIS is a pre-requisite of the GIS Specialist job, and additional guidance is provided in the Mapper-Lister-GIS manual.

Prepare Map Atlas for interviews

The Planning Team should work with the GIS Team to prepare an atlas of maps for interviewers. Detailed instructions to prepare the Atlas-Interview are provided in the Mapper-Lister-GIS manual. Each atlas map should be printed on A3 or A4 paper in coloured ink with the following elements for each PSU:

- Sampling area name
- PSU ID
- Latitude-longitude coordinate inside of PSU
- Atlas version description
- Scale bar
- North arrow

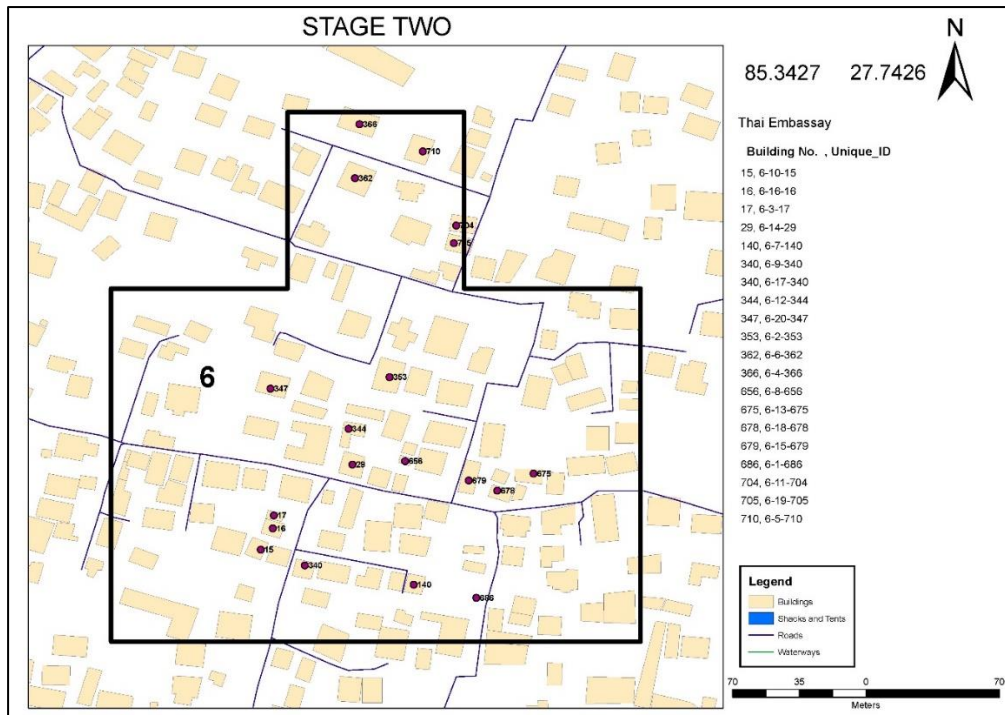


Figure 25. Paper field map for Interview phase in example PSU

Prepare tablets for interviews

Household and dwelling listing data collected during the mapping-listing phase should be preloaded on tablets for the interview phase. The process of tablet preparation for interviews is slightly different in one-stage and two-stage PSUs.

Two-stage PSUs:

- **Programme the questionnaire form** in Excel and upload it to the server following instructions in "How To: Setup GeoODK form in Excel" in Appendix C. Note that these are the same steps used to create the listing form. Ensure that the first fields in the questionnaire are the same as the household listing form.
- Compile a **.csv file of household listing data** for only those households selected for the survey (e.g. the 20 households selected per PSU).
- **Pull the .csv of household listing data into the questionnaire form** following instructions in "How To: Pull listing data into questionnaire form" in Appendix C. This will pre-populate such fields as ward, building ID, household number, and head of household name, making it easy to navigate to the correct household and saving the interviewer data entry time.
- **Upload the questionnaire to tablets** following instructions in "How To: Setup tablets" in Appendix C.
- See Appendix C "How To: Complete two-stage interview questionnaire" to **understand the one-stage interview protocol for interviewers**

One-stage PSUs:

- The paper maps are important for relocating the PSU and identifying all previously listed dwellings. Note that in Nepal, we pre-listed dwellings to be consistent with two-stage protocols, however, **other one-stage surveys might skip the dwelling listing phase, and directly list and interview households.**
- **Programme the questionnaire form in Excel** and upload it to the server following instructions in "How To: Setup GeoODK form in Excel" in Appendix C. Note that these are the same steps used to create the listing form. Ensure that the first fields in the questionnaire are the same as the household listing form. Key questions:
 - (a) **PSU ID:** A list of listed PSU IDs are pre-programmed
 - (b) Is the **dwelling already listed** in the dwelling listing form?: Yes/No
 - If No: A blank dwelling/household listing form is started
 - If Yes: A list of already-listed Dwelling IDs (in the selected post-field segment) are pre-programmed
 - (c) Note, a household listing should be completed for each dwelling before interview questions begin.
- **Upload the questionnaire to tablets** following instructions in "How To: Setup tablets" in Appendix C.
- See Appendix C "How To: Complete one-stage interview questionnaire" to **understand the one-stage interview protocol for interviewers**

Using PSU maps & dwelling/household listing for interviews

This manual does NOT provide detailed protocols to conduct household survey interviews. The interviewer manual should be consulted for these protocols.

However, the mapping-listing process is foundational to successful interviews. The final output of the mapping-listing process is a geographically accurate paper map of building locations and types in each PSU, and a complete list of dwellings (one-stage) or households (two-stage).

Replacing households (two-stage only). The majority of households listed during field mapping-listing were present at the time of interview. However, if the listed households had moved, we replaced the household in the following way:

1. Previously listed household had moved to another place and the dwelling was empty – interviewed a neighbouring household.
2. Previously listed household had moved and a new household was living in the dwelling – interview the current household.

We did NOT replace the household if the household refused to participate, or we did not find a household member after 3 visits.

Sample Weights

Sample weights are necessary to make accurate estimates in the population from a complex household survey. Sample weights are presented below. Inputs include three sources of information:

- GridSample: PSU population and strata population
- Fieldwork: Number of segments created in the PSU during pre-field enumeration, number of segments created in the PSU during post-field enumeration (one-stage only), total number of dwellings/households listed in the final (segmented) PSU, and number of households selected for sampling (two-stage only)
- Interview: Number of non-responding households per PSU (household questionnaire), and number of non-responding individuals per PSU (individual questionnaire)

Household sample (base) weight. The formulas use 2 indices: 1...*k* strata (or entire coverage area) and 1...*i* PSU. The household sample (design) weight – the probability that PSU *i* is selected – is given by:

$$w_{hh.b} = \frac{G_k/g_{ik}}{n_k} \times \frac{M_{ik}}{m_{ik}} \times (b1_{ik} * b2_{ik})$$

Where:

n_k is the number of selected PSUs in stratum *k*

G_k is the estimated total population in stratum *k* from GridSample

g_{ik} is the estimated population in PSU *i* in stratum *k* from GridSample

m_{ik} is the number of households sampled in PSU *i* and stratum *k* during fieldwork

M_{ik} is the number of total households enumerated in PSU *i* and stratum *k* during fieldwork

$b1_{ik}$ is the number of pre-field segments

$b2_{ik}$ is the number of post-field segments

Household response weight. Interviewers will list households and record household and individual response rates during fieldwork. After interviews are completed, calculate household response weight - the probability that cluster *i* is found and sampled, and households are found and respond – is given by:

$$w_{hh.r} = \frac{n_k}{n_{k*}} \times \frac{m_k}{m_{k*}}$$

Where:

n_k is the number of selected PSUs in stratum k

m_k is the number of households sampled in stratum k during fieldwork

n_{k^*} is the number of found and sampled PSUs in stratum k

m_{k^*} is the number of found and responded households in stratum k

Household sample weight. To calculate the raw household sample weight, multiply the sample base weight and household response weight like this:

$$w_{hh} = \frac{G_k/g_{ik}}{n_k} \times \frac{M_{ik}}{m_{ik}} \times (b1_{ik} * b2_{ik}) \times \frac{n_k}{n_{k^*}} \times \frac{m_k}{m_{k^*}}$$

Note that in one-stage samples, $\frac{M_{ik}}{m_{ik}}$ is equal to 1, and in two-stage samples, $b2_{ik}$ is equal to 1.

Individual sample weight. The individual sample weight includes four additional terms to account for the sampling of one adult among all eligible adults in the household, and the response rate of those adults. The individual sample weight is given by:

$$w_{ind.s} = \frac{G_k/g_{ik}}{n_k} \times \frac{M_{ik}}{m_{ik}} \times (b1_{ik} * b2_{ik}) \times \frac{n_k}{n_{k^*}} \times \frac{m_k}{m_{k^*}} \times \frac{U_{ik}}{u_{ik}} \times \frac{u_k}{u_{k^*}}$$

Where:

n_k is the number of selected PSUs in stratum k

G_k is the estimated total population in stratum k from GridSample

g_{ik} is the estimated population in PSU i in stratum k from GridSample

m_{ik} is the number of households sampled in PSU i and stratum k

M_{ik} is the number of total households enumerated in PSU i and stratum k

$b1_{ik}$ is the number of pre-field segments

$b2_{ik}$ is the number of post-field segments

n_{k^*} is the number of found and sampled PSUs in stratum k

m_k is the number of households sampled in stratum k during fieldwork

m_{k^*} is the number of found and responded households in stratum k

U_{ik} is the number of eligible adults in PSU i and stratum k

u_{ik} is the number of sampled adults in PSU i and stratum k

u_k is the number of sampled adults in stratum k

u_{k^*} is the number of responded adults in stratum k

Calculate sample weights in Excel. Calculate sample weights in Excel, or a statistical software programme such as Stata. See Figure 25 for example.

Normalizing sample weights. Household surveys are often “normalized” or “standardized” such that the sum of weighted respondents equals the sum of respondents. Conceptually, each observation in the sample represents slightly more or slightly less than 1 household or person. To normalize sample weights, apply the below formulas:

$$w_{hh_norm} = w_{hh} \times \frac{\sum(m_{ik*})}{\sum(w_{hh} \times m_{ik*})}$$

Where:

m_{ik*} is the number of adults with a completed interview in PSU i in stratum k
 w_{hh} is the raw household sample weight in PSU i in stratum k

$$w_{ind_norm} = w_{ind} \times \frac{\sum(u_{ik*})}{\sum(w_{ind} \times u_{ik*})}$$

Where:

u_{ik*} is the number of adults with a completed interview in PSU i in stratum k
 w_{ind} is the raw individual sample weight in PSU i in stratum k

PSU ID	Stratum ID	g _{ik}	n _k	G _k	1/b _{ik}	M _{ik}	m _{ik}	n _k *	m _k	m _k *	m _{ik} *	w _{hh}	w _{hh} *m _{ik} *	w _{hh} norm
1	1	872.0	15	112020.6	0.071	12	12	15	292	292	12	119.901	1438.809	0.930
2	2	91.4	5	1318.6	0.059	11	11	5	95	90	11	51.763	569.392	0.401
3	1	830.1	15	112020.6	0.063	24	24	15	292	292	24	143.936	3454.476	1.116
4	1	245.5	15	112020.6	0.091	24	24	15	292	292	24	334.641	8031.387	2.595
6	1	888.2	15	112020.6	0.063	16	16	15	292	292	16	134.527	2152.435	1.043
7	2	230.7	5	1318.6	0.071	27	27	5	95	90	26	16.890	439.150	0.131
8	2	94.1	5	1318.6	0.024	12	12	5	95	90	12	124.305	1491.654	0.964
9	1	1142.9	15	112020.6	0.067	25	25	15	292	292	25	98.016	2450.404	0.760
10	1	317.1	15	112020.6	0.077	16	16	15	292	292	16	306.135	4898.162	2.374
12	1	702.8	15	112020.6	0.083	18	18	15	292	292	18	127.509	2295.163	0.989
13	1	617.3	15	112020.6	0.071	7	7	15	292	292	7	169.376	1185.629	1.314
14	2	172.7	5	1318.6	0.091	15	15	5	95	90	15	17.735	266.018	0.138
15	2	133.7	5	1318.6	0.143	30	30	5	95	90	26	14.573	378.888	0.113
16	1	721.4	15	112020.6	0.053	21	21	15	292	292	21	196.692	4130.526	1.525
18	1	728.6	15	112020.6	0.111	25	25	15	292	292	25	92.244	2306.104	0.715
20	1	844.2	15	112020.6	0.059	19	19	15	292	292	19	150.396	2857.520	1.166
22	1	871.0	15	112020.6	0.067	12	12	15	292	292	12	128.607	1543.286	0.997
23	1	840.8	15	112020.6	0.063	15	15	15	292	292	15	142.106	2131.588	1.102
24	1	605.1	15	112020.6	0.111	35	35	15	292	292	35	111.083	3887.898	0.861
25	1	666.7	15	112020.6	0.077	23	23	15	292	292	23	145.617	3349.183	1.129
TOTAL						387	387			382	2626.051	49257.672		
										Σm_{ik}^*	$\Sigma(w_{hh} * m_{ik}^*)$			

PSU ID	Stratum ID	w _{hh}	u _k	U _{ik}	u _{ik}	u _{ik} *	u _k *	w _{ind}	w _{ind} *u _{ik} *	ind norm	
1	1	119.901	292	34	12	12	288	344.437	4133.245	0.979	
2	2	51.763	90	35	11	11	88	168.443	1852.878	0.479	
3	1	143.936	292	67	24	23	288	407.404	9370.283	1.158	
4	1	334.641	292	62	24	24	288	876.496	21035.913	2.491	
6	1	134.527	292	32	16	16	288	272.791	4364.659	0.775	
7	2	16.890	90	84	26	26	88	55.809	1451.037	0.159	
8	2	124.305	90	30	12	12	88	317.824	3813.889	0.903	
9	1	98.016	292	63	25	25	288	250.431	6260.782	0.712	
10	1	306.135	292	40	16	15	288	775.968	11639.513	2.205	
12	1	127.509	292	49	18	18	288	351.929	6334.722	1.000	
13	1	169.376	292	19	7	7	288	466.119	3262.832	1.325	
14	2	17.735	90	39	15	14	88	47.158	660.208	0.134	
15	2	14.573	90	54	26	25	88	30.954	773.852	0.088	
16	1	196.692	292	52	21	21	288	493.811	10370.024	1.403	
18	1	92.244	292	86	25	25	288	321.727	8043.178	0.914	
20	1	150.396	292	46	19	19	288	369.173	7014.293	1.049	
22	1	128.607	292	27	12	12	288	293.385	3520.620	0.834	
23	1	142.106	292	43	15	15	288	413.028	6195.421	1.174	
24	1	111.083	292	136	35	34	288	437.631	14879.453	1.244	
25	1	145.617	292	52	23	22	288	333.793	7343.441	0.949	
TOTAL				1050	382	376	4760	7028.311	132320.242		
						Σu_{ik}^*			$\Sigma(w_{ind} * u_{ik}^*)$		

Figure 26. Example household and individual sample weights calculations

Review

Answer the questions below. To see if your answer is correct, go to Appendix D.

1. Sample weights include _____, and thus are calculated after _____.
 - a. Vacated buildings, fieldwork
 - b. Separate estimates for each segment, post-field segmentation
 - c. Response rates, interviews

2. Post-field segmentation is performed in:
 - a. One-stage PSUs only
 - b. Two-stage PSUs only
 - c. A and B

References

Campbell (2017). Working with people and communities in urban humanitarian crises. <http://www.urban-response.org/resource/25143>.

DHS (2012). Sampling and household listing manual. https://dhsprogram.com/pubs/pdf/DHSM4/DHS6_Sampling_Manual_Sept2012_DHSM4.pdf.

Eley (2017). Addressing Inequities in Urban Health: Do Decision-Makers Have the Data They Need? <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4899330/>.

Thomson (2012). Modelling strategic interventions in a population with a total fertility rate of 8.3. <https://bmcpublikealth.biomedcentral.com/articles/10.1186/1471-2458-12-959>

Thomson (2017). GridSample: an R package to generate household survey primary sampling units (PSUs) from gridded population data. <https://ij-healthgeographics.biomedcentral.com/articles/10.1186/s12942-017-0098-4>.

Appendix A: Exercises

Exercise: Reproduce Nepal SUE survey sample size calculation

Step 1: Download the WHO-based Excel Sample Size & Budget Calculator from the [GridSample website](#).

Step 2: Read the following description of the SUE Nepal survey sample size calculation prepared by Joe Hicks (University of Leeds).

“The sample will consist of 1200 household interviews in 60 sampling areas of 20 households each. This will enable us to generate sufficient data to assess the extent to which our approach identifies the urban poor, and to estimate key mental health and injury indicators with a maximum margin of error of +/- 4.27% with 95% confidence (assuming the most conservative scenario where an indicator is estimated at 50%). This assumes a design effect of 1.41 (the mean DEFT across all indicators for men and women in urban areas in Nepal DHS 2011), a household and an individual response rate of 0.98 and 0.93 respectively (based on conservative estimates from response rates in urban areas in Nepal 2011 DHS), and one eligible individual per household.”

Step 3: The tool provides sample size calculates for Classification, Estimation, and Comparison surveys.

Question: For which type of survey are we calculating sample size?

Step 4: Enter the above parameters into the correct fields.

Question: What values did you enter/calculate in the following fields?

(A) Number of strata (N_strata)	
(B) Effective sample size (ESS)	
(C) Design effect (DEFF)	
(D) Average number of households to visit to find 1 eligible respondent	
(E) Non-response inflation factor	

Total number of households to visit (household sample size)	
Number of clusters per stratum	
Number of households per cluster	

Step 5: It is common for survey designers to round the sample size.

Question: Based on your sample size calculations, what sample sizes do you recommend for this multi-stage cluster survey of Kathmandu Valley? Did you agree with Joe?

Total households	
Total clusters	
Households per cluster	

Answers:

Q3: Estimation

Q4: $N_{\text{strata}}=1$, $ESS=547$, $DEFF=1.9881$, $m=20$, households to find 1 eligible respondent=1, non-response inflation factor=1.097 ($0.98*0.93=0.9114$). Households to visit=1194, clusters per stratum=55, households per cluster=22.

Alternatively, you may have played around with m to make the final "households per cluster"=20. If so, you entered $N_{\text{strata}}=1$, $ESS=547$, $DEFF=1.9881$, $m=18$, households to find 1 eligible respondent=1, non-response inflation factor=1.097 ($0.98*0.93=0.9114$). Households to visit=1194, clusters per stratum=61, households per cluster=20.

Q5: Total households=1200, Total clusters=60, Households per cluster=20.

Exercise: Become familiar with the GridSample website

Step 1: Open an internet browser, and navigate to gridsample.org.

Step 2: Navigate to the FAQ page, and read the section “Why use GridSample?”

Question: Of the three reasons that people use GridSample, which two reasons are most relevant to the SUE surveys?

Step 3: Take a few minutes to review another question on the FAQ page.

Question: What is one thing you learned that might be helpful to share with others?

Step 4. Navigate to the Tutorial page.

Question: What are the four approaches available to implement gridded samples with GridSample output?

1. _____

2. _____

3. _____

4. _____

Step 5. Navigate to the Tool page. We will now enter the parameters for a one-stage sample in Dhaka, Bangladesh.

- Sample name: "Test – One-stage Dhaka"
- Country: Bangladesh
- Admin 2: Dhaka
- WorldPop Dataset: BGD_ppp_v2b_2020_UNadj
- Stages: One-stage cluster sample
- Stratification: No
- Spatial representation: No
- Target population: "Household"
- Number of target population members per household: 1
- Average household size: 4.5
- Number of PSUs per stratum: 10
- Number of households sampled per urban/rural PSU: 20
- Number of households sampled per stratum: 200
- Input grid cells are: 1 (100m X 100m)
- Exclude cells fewer than: 0 population
- Do not generate PSUs larger than: Infinity
- Random number: 2017
- Growth algorithm: on

Answers:

Q2: Outdated or inaccurate census sample frame, Small primary sampling units (PSUs)

Q4: Gridded PSUs, manual PSUs, non-probability (e.g. random walk), and simple random sample of households

Exercise: Sample frame scale

1. Open www.gridsample.org in your web browser, and navigate to the Tool.
2. Generate two new samples using the following parameters. We provide this example for Dhaka; adopt using admin areas in your own city. The only differences between the samples are the sample name, and size of the input grid cells:

- Email: [Enter an email address that you check regularly]
- Sample name: "SUE Dhaka scale test 100m" in the first sample, and "SUE Dhaka scale test 200m" in the second sample
- Country: Bangladesh
- Admin 2: Dhaka
- WorldPop Dataset: BGD_ppp_v2b_2020_UNadj
- Stages: One-stage cluster sample
- Stratification: No
- Spatial representation: No
- Target population: "Household"
- Number of target population members per household: 1
- Average household size: 4.5
- Number of PSUs per stratum: 10
- Number of households sampled per urban/rural PSU: 20
- Number of households sampled per stratum: 200
- Input grid cells are: 1 (100m X 100m)
- Exclude cells fewer than: 0 population
- Do not generate PSUs larger than: Infinity
- Random number: 2017
- Growth algorithm: on

3. GridSample will email your selected PSUs as a shapefile and a KMZ file. Complete the "Sample frame scale worksheet" below, following these instructions.

Remember, we are still at the beginning of the sample workflow. This is what we know:

- Some buildings are not residential
- Buildings might have multiple levels
- There may be multiple dwellings on each level
- There may be multiple households in each dwelling

This is what we do not know:

- A complete picture of building locations (in OpenStreetMap)
- How many dwellings or households are in each PSU

So, we are going to make an educated guess based on Google Earth imagery, including apparent height of buildings.

Step 1: Copy the WorldPop population estimate for each PSU into the form.

After submitting your GridSample job, the user receives an email with a link to download the sample PSU boundaries. The download includes a shapefile (comprised of multiple subfiles including a dbf) which can be opened in ArcGIS or excel, and a kml file which can be opened in Google Earth.

Download and unzip this file, then open the dbf in Excel. Copy the PSU population estimates (variable name: psu_pop) into the worksheet, and round population to whole numbers.

Step 2: Use the average household size (entered in How To: Use GridSample exercise) to estimate the number of households per PSU. Note:

PSU households = PSU population ÷ average household size

Step 3: Open the kml files in Google Earth, and estimate the number of single family and multi-family buildings.

You may need to download Google Earth, first, from:
www.google.com/earth/download/gep/agree.html

Once Google Earth is installed on your computer, you only need to double click on the kml file to open it in Google Earth.

Step 4: Look closely at the imagery for each PSU, and answer the following question by circling YES or NO. Does Google Earth appear to show a similar # of HHs as WorldPop? You can answer NO if the imagery appears to have far more or far fewer households than WorldPop.

Step 5. Looking at Google Earth imagery, answer the following question by circling YES or NO: Do you think there are 20+ households in this PSU?

Step 6 (Nepal only). In Nepal, this exercise was repeated for two-stage PSUs with the target of 200 households each. Looking at Google Earth imagery, answer the following question by circling YES or NO: Do you think there are 100+ households in this PSU?

P S U #	Steps 1 and 2. WorldPop	Step 3: Google Earth and OpenStreetMap	Step 4. Does Google Earth show a similar # of HHs as WorldPop?	Step 5. Do you think there are 20+ households in this PSU?	Nepal only. Do you think there are 100+ households in this PSU?
First sample with 100m X 100m grid cells					
1	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
2	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
3	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
4	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
5	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
6	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
7	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
8	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
9	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
10	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
Second sample with 200m X 200m grid cells					
1	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
2	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
3	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
4	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
5	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
6	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
7	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
8	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
9	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO
10	Population: _____ Households: _____	# singlefam blgs: ____ # multifam blgs: ____	YES / NO	YES / NO	YES / NO

4. This guidance is subject to change as we learn more about gridded population sampling, but here are our initial recommendations. We are aware that this advice might conflict; use your best judgement to balance feasibility and accuracy when selecting a sample frame scale.
- Feasibility/accuracy: If you circled 9-10 YESs in step 4, and 9-10 YESs in step 5, then this sample frame scale is appropriate for one-stage sampling
 - Feasibility/accuracy:: If you circled 9-10 YESs in step 4, and 9-10 YESs in step 6 (Nepal only), then this sample frame scale is appropriate for two-stage sampling
 - Feasibility: If you expect to find more than 100 households in 4+ PSUs of the 200m X 200m sample frame, then use a 100m X 100m sample frame
 - Accuracy: If you circled NO in step 4 in 4+ PSUs of the 100m X 100m sample frame, then use a 200m X 200m sample frame. (Although this might result in far more households than desired, it may be preferable to aggregate the sample frame to 200m X 200m cells to improve accuracy and then manually segment PSUs, than to sample from an inaccurate sample frame.)

Appendix B: Forms

Training agenda

**Surveys for Urban Equity
HERD Int/Leeds University
SUE Survey Planning Team Training
14 – 15 Aug 2017**

Date: 14 August **OFFICE WORK** **Venue: HERD International**

Time	Topic
9:00am – 9:15am	Welcome, introductions
9:15am – 10:15am	Overview, definitions
10:15am – 10:45am	Workflows
10:45am – 11:00am	Tea break
11:00am – 12:00pm	GridSample
12:00pm – 1:00pm	Lunch break
1:00pm – 2:00pm	Sample frame
2:00pm – 2:15pm	Select samples
2:15pm – 3:00pm	Office enumeration
3:00pm – 3:20pm	Tea break
3:20pm – 3:40pm	Assign PSUs
3:40pm – 4:10pm	GIS, IT summary
4:10pm – 4:45pm	Jobs, schedule
4:45pm – 5:00pm	Debrief

Date: 15 August **FIELD WORK** **Venue: HERD International**

Time	Topic
9:00am – 9:45am	Questions
9:45am – 10:15am	Jobs, schedule
10:15am – 10:30am	Tea break
10:30am – 11:15am	Field enumeration overview
11:15am – 12:00am	Field enumeration setup
12:00pm – 1:00pm	Lunch break
1:00pm – 1:45pm	Listing overview
1:45pm – 2:30pm	Listing setup
2:30pm – 3:00pm	Interview discussion
3:00pm – 3:20pm	Tea break
3:20pm – 4:00pm	Session 1: Consults with Dana, HERD Team
4:00pm – 4:40pm	Session 2: Consults with Dana, HERD Team
4:40pm – 5:00pm	Debrief

Participants (email addresses removed for privacy)

HERD International:

- Radheshyam Bhattarai
- Rajeev Dhungel
- Subash Gajurel
- Sudeepa Khanal
- Shraddha Manandhar

CIPRB:

- Silvia Junnatul Ferdoush

ARK Foundation:

- Tarana Ferdous

Hanoi School of Public Health:

- Duong Minh Duc
- Nguyen Bao Ngoc

Leeds University:

- Ak Narayan Poudel

Southampton University:

- Dana Thomson

Mapper-Lister Job Announcement

Role: Digital Mapping and Household Listing Specialist

Duration: Full time, 16 Aug – 12 Sep 2017

Salary: [Complete]

About HERD International

HERD International is a national organisation with a vision of achieving health for all through strengthened health systems. We believe strongly in evidence-based actions and focus our expertise on generating evidence, and effectively linking it to research, policy, and most importantly, practice.

Position background

This position will be with the SUE research project – “Surveys for Urban Equity: getting, and using, the data to respond to non-communicable diseases in urban areas”. A key activity in this project is a representative household survey about health issues that affect people in the Kathmandu Valley. 10-12 Digital Enumeration and Household Listing Specialists will play an essential role in the survey by mapping all of the areas selected for the survey, and listing all of the households in those areas. After your work, 1200 households will be selected and interviewed.

The SUE survey will pilot several new technologies for household selection, enumeration, and listing. We will provide classroom and field training on these tablet-based technologies. If hired, we hope that you will participate in an optional focus group discussion at the end of the work to share your experiences and recommendations about these cutting-edge methods. There are many challenges to accurately enumerating and listing households in cities, so survey teams around the world are interested in the methods to be piloted in this survey, and in your reflections to implement them.

Duties

- Fully participate in a three-day training, and pass a practical and a written evaluation
- Travel independently to survey areas across the Kathmandu Valley, bus fare provided
- Update OpenStreetMap roads and building using a desktop computer
- Update OpenStreetMap roads, buildings, and points of interest using tablets
- List all dwellings and households in the survey areas using a digital form on tablets
- Speak with residents and business owners to generate a complete household listing, including of households that not present at the time of the field visit

- Answer questions and distribute printed information about the survey to community members, if asked
- Duties will be performed in teams of 2 during daylight hours, with remote support of a supervisor
- Submit daily work reports
- Attend regular in-person meetings at the HERD International office

Profile requirements

Required

- Ability to work independently, and ask appropriate questions as needed
- Able to read a map
- Basic knowledge of GPS devices
- Comfortable using mobile technologies, particularly for data collection or navigation
- Bachelor's degree, or relevant field work experience
- Language: Nepali

A plus

- Experience using a geographic information system (for example QGIS or ArcGIS)
- Experience collecting household survey data by paper or tablet
- Experience cleaning or managing household survey data
- Language: English
- Genuine interest in supporting policy-makers and decision-making with high quality, accurate data

How to apply

- Submit a CV to _____ by [DATE]

GIS Job Announcement

Role: GIS Specialist

Duration: Full time, 16 Aug – 12 Sep 2017

Salary: [Complete]

Number of positions: 1

About HERD International

HERD International is a national organisation with a vision of achieving health for all through strengthened health systems. We believe strongly in evidence-based actions and focus our expertise on generating evidence, and effectively linking it to research, policy, and most importantly, practice.

Position background

This position will be with the SUE research project – “Surveys for Urban Equity: getting, and using, the data to respond to non-communicable diseases in urban areas”. A key activity in this project is a representative household survey about health issues that affect people in the Kathmandu Valley. The GIS specialist is a key member on the team, providing spatial data management and map production throughout the survey fieldwork.

The SUE survey will pilot several new technologies for household selection, enumeration, and listing. You will join classroom and field training with the field data collection team so you understand the innovative tablet-based technologies that are being used in the SUE surveys. If hired, we hope that you will participate in an optional focus group discussion at the end of the work to share your experiences and recommendations about these cutting-edge methods. There are many challenges to accurately enumerating and listing households in cities, so survey teams around the world are interested in the methods to be piloted in this survey, and in your reflections to implement them.

Duties

- Fully participate in a three-day training
- Generated dozens of field maps using an Atlas tool
- Digitize sampling area (polygon) boundaries based on satellite imagery
- Convert between spatial data formats (shapefile, GPX, MBTiles)
- Generate map reports to support the planning team with fieldwork quality checks
- Attend regular in-person meetings at the HERD International office
- Closely follow provided training materials – and look up additional materials as needed - to perform GIS tasks with minimal supervision

Profile requirements

Required

- Intermediate-to-advanced skills in ArcGIS or QGIS – ideally demonstrated with printed or web-based GIS work linked to your CV
- 1 year of classroom-based GIS training
- 6+ months of GIS working experience
- Experience using GPS devices
- Language: Nepali

A plus

- Experience cleaning or managing household survey data
- Bachelor's or higher degree
- Language: English
- Genuine interest in supporting policy-makers and decision-making with high quality, accurate data

How to apply

- Submit a CV to _____ by [DATE]

Planning team OFFICE tracking form

Tracking sheet: Plan

PSU ID	Date of office enumeration	Assignment #	Date of one-stage boundaries	Date of pre-field segmentation	Updated Assignment #	Date of Atlas-E-L	Date final paper map in OneDrive
Responsible	Enumerator-Lister, GIS	Planning	GIS	GIS	Planning	GIS	Enumerator-Lister
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							
36							
37							
38							
39							
40							
41							
42							
43							

Planning team FIELD tracking form

Tracking sheet: I

PSU ID	Updated assignment # (Type)	Stages (One or two)	Team name	Phone number 1	Phone number 2	District name	Area name
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							
36							
37							
38							
39							
40							
41							
42							
43							
44							

Office mapping tracking form



Tracking sheet: Office enumeration

ID	Date	Enumerator name	Time taken	Informal housing (yes/no)	Date checked by HERD Int	Remarks
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						
27						
28						
29						
30						
31						
32						
33						
34						
35						
36						
37						
38						
39						
40						
41						
42						
43						
44						
45						
46						
47						
48						
49						
50						
51						
52						
53						
54						
55						
56						
57						
58						
59						
60						
61						
62						
63						
64						
65						
66						
67						
68						
69						
70						
71						
72						
73						
74						
75						
76						
77						
78						

Tracking Sheet: GIS Team



Tracking sheet: GIS

ID	GIS + Planning Team		GIS Team		GIS + Planning Team		GIS Team		GIS + Planning Team		GIS Team		GIS + Planning Team		GIS Team	
	Review FSUs assignment	Assignment Non-perm. housing	One-stage boundaries - Nepal only	Pre-field segmentation	Assignment	Create GPX & MBTiles	Atlas E-L	Download permanent OSM data	Digitize shack/tenement-residential	Atlas Interviewer	One-stage post-field seg	One-stage post-field seg	Download permanent OSM data	Digitize shack/tenement-residential	Atlas Interviewer	
	(Date)	(type - 1)	(Date)	seg (y/n)	(type - 2)	(Date)	(Date)	seg (y/n)	(Date)	seg (y/n)	(Date)	(Date)	seg (y/n)	(Date)	seg (y/n)	
1																
2																
3																
4																
5																
6																
7																
8																
9																
10																
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																
21																
22																
23																
24																
25																
26																
27																
28																
29																
30																
31																
32																
33																
34																
35																
36																
37																
38																
39																
40																
41																
42																

Tracking sheet: GIS

ID	GIS + Planning Team Review PSU's assignment		GIS Team One-stage boundaries - Nepal only	GIS + Planning Team Pre-field segmentation		GIS Team Create GPX & MBTiles	GIS Team Atlas E-L	GIS + Planning Team One-stage post-field seg		GIS Team Download permanent OSM data	GIS Team Digitize shack/tent-residential	GIS Team Atlas Interviewer
	(Date)	(type - 1) (v/n)		(Date)	seg (v/n) (type - 2)			(Date)	seg (v/n)			
43												
44												
45												
46												
47												
48												
49												
50												
51												
52												
53												
54												
55												
56												
57												
58												
59												
60												
61												
62												
63												
64												
65												
66												
67												
68												
69												
70												
71												
72												
73												
74												
75												
76												
77												
78												
Plan. Team Initials												
Total Time												

Tracking Sheet: Mapping-Listing Team



Team Swan
Team member names:

Tracking sheet: Field enumeration

Day	PSU ID	Stages	Task	Name	Date	Time started	Time ended	Total time	Date checked by HERD Int	Remarks (completed/# of buildings, roads edited/issues)
			Field paper							
			Listing							
			ID Editor							
			Field paper							
			Listing							
			ID Editor							
			Field paper							
			Listing							
			ID Editor							
			Field paper							
			Listing							
			ID Editor							
			Field paper							
			Listing							
			ID Editor							
			Field paper							
			Listing							
			ID Editor							
			Field paper							
			Listing							
			ID Editor							
			Field paper							
			Listing							
			ID Editor							
			Field paper							
			Listing							
			ID Editor							
			Field paper							
			Listing							
			ID Editor							
			Field paper							
			Listing							
			ID Editor							

Detailed mapping-listing schedule

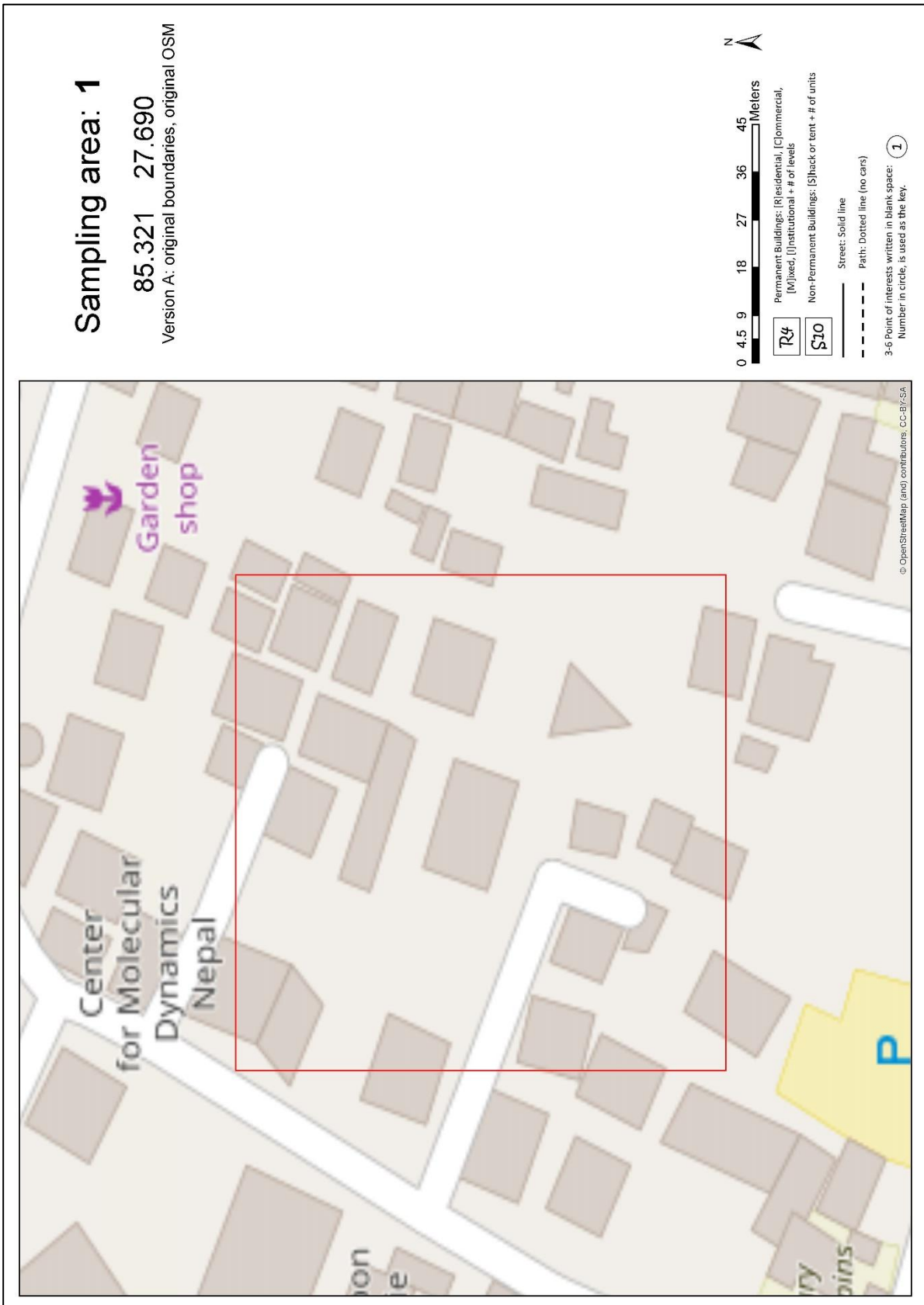
Enumeration schedule

Day	Team Swan	Team Giraffe	Team Tiger	Team Eagle	Team Kangaroo	Team Dolphin	GIS
16-Aug	Train	Train	Train	Train	Train	Train	Train
17-Aug	Train	Train	Train	Train	Train	Train	Train
18-Aug	Train	Train	Train	Train	Train	Train	Train
19-Aug	Off	Off	Off	Off	Off	Off	Off
20-Aug	Off	Off	Off	Off	Off	Off	Off
21-Aug	iD Editor	iD Editor	iD Editor	iD Editor	iD Editor	iD Editor	iD Editor
22-Aug	iD Editor	iD Editor	iD Editor	iD Editor	iD Editor	iD Editor	iD Editor
23-Aug	iD Editor	iD Editor	iD Editor	iD Editor	iD Editor	iD Editor	iD Editor
24-Aug	Train	Train	Train	Train	Train	Train	Segment, Atlas
25-Aug	Off	Off	Off	Off	Off	Off	
26-Aug	Off	Off	Off	Off	Off	Off	
27-Aug	Pick-up	Pick-up	Pick-up	Pick-up	Pick-up	Pick-up	
28-Aug							
29-Aug							
30-Aug	31,1	36,6	41,11	46,16	51,21	56,26	
31-Aug							
01-Sep							
02-Sep	32,2	37,7	42,12	47,17	52,22	57,27	
03-Sep							
04-Sep							
05-Sep	33,3	38,8	43,13	48,18	53,23	58,28	
06-Sep	Off	Off	Off	Off	Off	Off	
07-Sep							
08-Sep							
09-Sep	34,4	39,9	44,14	49,19	54,24	59,29	
10-Sep							
11-Sep							
12-Sep	35,5	40,10	45,15	50,20	55,25	60,30	
13-Sep	Backup, Group	Backup, Group	Backup, Group	Backup, Group	Backup, Group	Backup, Group	Segment
14-Sep	Focus	Focus	Focus	Focus	Focus	Focus	Segment
15-Sep	Group	Group	Group	Group	Group	Group	Atlas maps

Note: In a 3 day period, enumerators will map and list households/dwelling in one 2-stage PSU (first 2 days) and one 1-stage PSU (morning of 3rd day). In addition, they will come to HERD Int office and update OSM (afternoon of 3rd day).

1-Stage	2-Stage
16	14
20	18
21	22
33	26
36	28
3	9
7	19
8	24
13	27
23	29
40	39
62	44
65	48
66	54
72	70
49	64
55	71
58	76
67	77
75	78
1	2
4	5
11	6
12	10
15	25
31	32
35	38
41	43
42	73
46	74

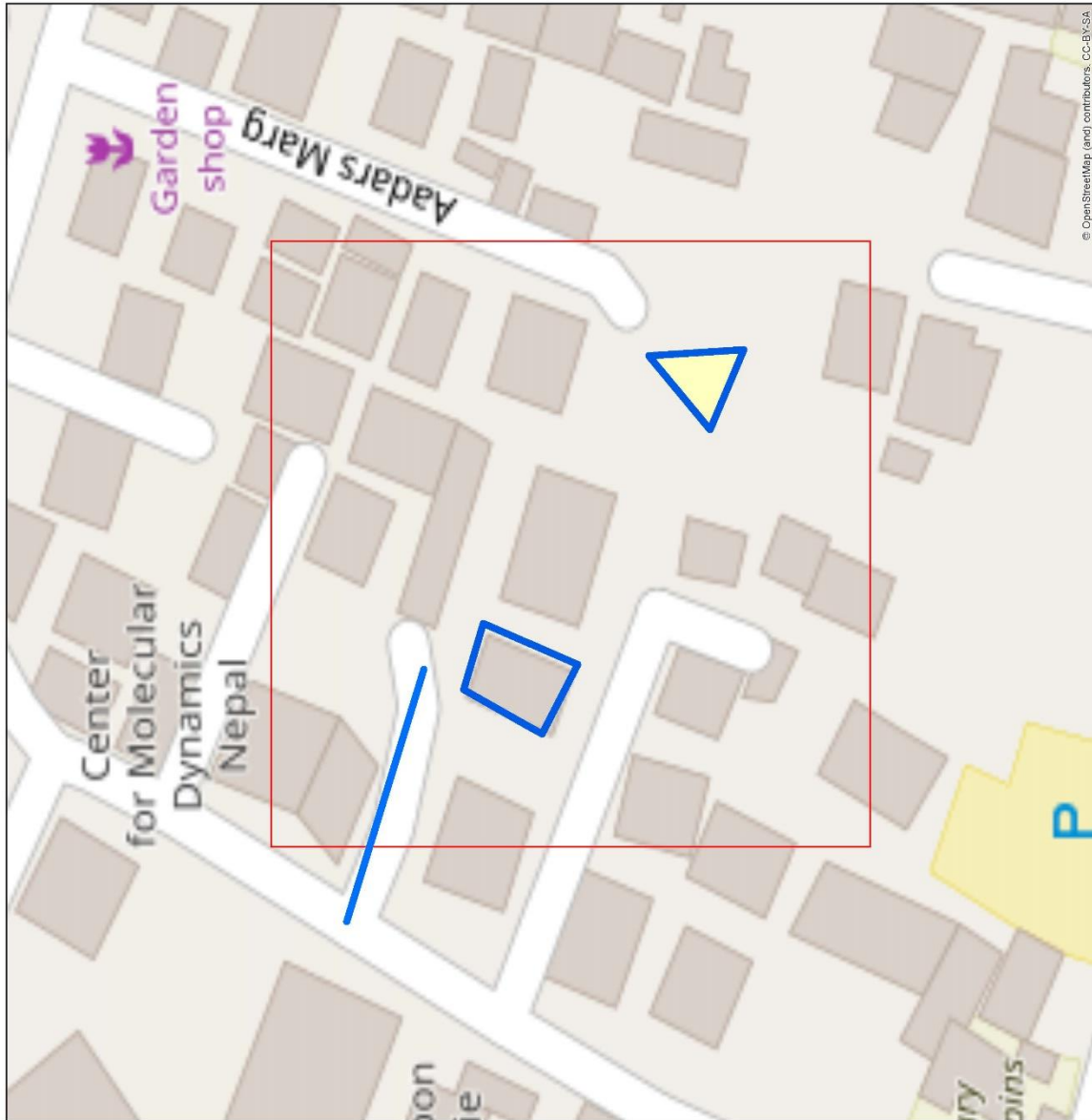
Example mapping-listing field exam map and answer key



Sampling area: 1

85.321 27.690

Version A: original boundaries, original OSM



Mapping-listing field exam score sheet

Prompts:

- While located at the office, ask the Mapper and Lister: "Please guide us to the sampling area."
- Once located at the sampling area, ask the Mapper and Lister: "What is your planned route??"

As questions arise, encourage the Mapper and Lister to discuss and arrive at an answer together.

Evaluator: _____

Mapper: _____

	0	1	2	Feedback / notes
Navigation				
Used MAPS.ME, OSMAnd, and/or printed map to arrive in sampling area				
Confirmed arrival in sampling are on printed map (using intersection, point of interest, building layout)				
Mapping - field				
Drew missing building on paper map				
Added building in correct location				
Added building correctly marked as R/C/M/I/S with number of levels				
Drew missing streets on paper map				
Added street was in correct location				
Correctly removed non-existent building on paper map				
Added two sensible "Points of Interest" (POI)				

0 – did not complete; **1** – completed partially, or with difficulty; **2** – completed normally

Observe the mapper's interaction with community members and comment on her/his greetings, presentation, politeness, and responses to questions:

Evaluator: _____

Lister: _____

	0	1	2	Feedback / notes
Listing				
Created a unique entry for building 1				
Correctly entered number of levels				
Correctly entered number of dwellings				
Correctly entered number of households				
Listed a head of household name for HH				
Created a unique entry for building 2				
Correctly entered number of levels				
Correctly entered number of dwellings				
Correctly entered number of households				
Listed a head of household name for HH				
Appropriately introduced self and survey				
Correctly discussed the definition of a "household" with residents and neighbours				

0 – did not complete; **1** – completed partially, or with difficulty; **2** – completed normally

Observe the lister's interaction with community members and comment on her/his greetings, presentation, politeness, and responses to questions:

Daily training feedback form

1-3 things you liked the most	1-3 things to be improved
New thing learned today	Thing which was not clear

1-3 things you liked the most	1-3 things to be improved
New thing learned today	Thing which was not clear

Appendix C: How-Tos for IT team

How To: Setup GoeODK server

Background

SMAP Server is a robust, flexible and standards based data collection tool. This data can be applied to the core systems that manage our business. SMAP server provide hosted servers including a free server to simplify the data collection process. Currently SMAP server development focus is on helping customers using these servers put their data to work.

1. Browse sg.smmap.com.au as shown in fig below.



2. It will open SMAP server page. SMAP Server is Open Source and setup file can be downloaded from [here](#) to create own server on premises. However please refer On Premises SMAP Server-Setup file attached herewith.



On Premises SMAP
Server-Setup.pdf

3. Hence SMAP server is Open Sources and hosted free SMAP server with URL sg.smmap.com.au which is designed organization based access system and isolated from each other.

Two approaches to use sg.smmap.com.au.

- a. 1st approach.
 - a. HERD International has its own page on sg.smmap.com.au server where we can create projects like
 - i. SUE-Nepal
 - ii. SUE-Vietnam
 - iii. SUE-Bangladesh
 - b. Under these project, we can upload multiple digital enumeration and survey tools.
- b. 2nd approach
 - a. Create own organization page isolated from HERD International.
 - b. Create multiple projects as per requirement.
 - c. Upload multiple digital enumeration and survey tools under created projects.

1st approach

- a. Request HERD International ICT team (Rajeev.dhungel@herdint.com) to create projects (ex: SUE-Vietnam or SUE-Bangladesh) and send list of users whom to provide access on this projects.
- b. While sending list of users, please assign role to users.
 - a. There are total three roles can be assigned to users.
 - i. Admin :- Can able to perform all task including
 1. Creating / Deletion / Modification of new users.
 2. Creation / Deletion / Modification of new projects.
 3. Able to assign role to created users.
 4. Assign created user access to projects.
 - ii. Analyst: - Can able to perform all task excluding creating / deleting / modification of users and projects. Unable to assign role to created users and give user access to projects.
 - iii. Enum:- Only allow to download forms to a phone and submit completed form from tablet to server.
 - b. Hence in this approach HERD international ICT team is admin so only two roles (Analyst and Enum) can be assigned.

2nd approach

- a. Browse URL sg.smap.com.au. Click on Register button.



- b. It open Registration Form

Registration Form

Organisation Name	<input type="text" value="HERD International"/>	✓
Your Name	<input type="text" value="Rajeev Dhungel"/>	✓
Your Email	<input type="text" value="rajeev.dhungel@herdint.com"/>	✓
Organisation Website	<input type="text" value="http://herdint.com"/>	✓

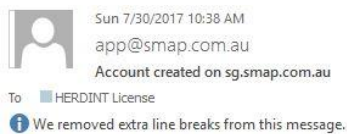
I agree to the terms and conditions

- i. Organization Name:- Type your organization name
- ii. Your Name:-Type full name of user who have admin access
- iii. Your Email: Type official email address of user whose name is placed under Your Name section.
- iv. Organization Website:- Type valid URL of organization.
- v. Tick I agree to the terms and condition and click submit button.

Registration accepted. An email has been sent to rajeev.dhungel@herd.org.np with a link that you can use to set your password.



c. SMAP server sends email with a link from email address (app@smap.com.au) to email address used during registration.



Rajeev Dhungel has given you access to a Smap server with address <https://sg.smap.com.au> You will need to specify your password before you can log on. To do this click on the following link <https://sg.smap.com.au/resetPassword.html?token=381a4ff9-0cd5-4cd0-8b20-e2e3220e855f>

Your user name is: rajeev.dhungel@herdint.com

The link is valid for 48 hours
Do not reply to this email address it is not monitored. If you don't think you should be receiving these then send an email to null.

Contains of email are

- a. User name with admin role
- b. Reset password to login SMAP server portal.

When click on given link it open SMAP reset password.


Smmap Reset Password

Type desired password and click submit.

Operation of SMAP Server.

Access to Profile

- a. At the top right of most pages we see a button with human picture. Click in the button to show your profile

Modules ▾ Monitoring Users Shared Resources Oversight Logs Help 

rajeevherd@HERD ✕

Name

Language

Email

Reset Password

A dialog is shown that allows you to

- Change your displayed name (not the identifier you logon with)
- Change the language you want the server to use
- Reset your password
- Logout

User Management

Note: - Only user with admin privilege can perform this tasks.

- b. To get to the user management page, Select the “Users” menu from inside the “admin” module
- c. Using this page admin can perform following actions.
 - i. Users
 - ii. Projects
 - iii. Roles
 - i. Requires Security Manager Privilege
 - ii. Restrict access to forms, columns in forms and rows depending on their data contents.
- d. Users – Create a new user, Editing an existing user

Users Projects

Show users with security group:

Show users in project:

Select	User Id	Name
<input type="checkbox"/>	test1	Test
<input type="checkbox"/>	thirdparty	thirdparty
<input type="checkbox"/>	umhm	umhm

Add a new user

Edit an existing user

e. It open user dialog box.

- i. Ident:- This is the identifier that the user will logon with .
- ii. Name:- This is the actual name of the user.
- iii. Email:- User email used for password resets.
- iv. Password:- There are two options for setting the password.
 - a. The user chooses their own password when after they are sent an email.
 - b. The administrator sets the password for the user. Setting the password directly is the best approach when creating user ids for anonymous data collectors. If you are creating an account for an analyst then using the automated email is probably better.
- v. Security Groups:-A user can belong to up to 3 security groups
 - a. Admin: can manage users and projects
 - b. Analyst: can view, or delete collected data
 - c. Enum: can download forms and submit results only (This is quite limited access)
- vi. Project:- Users will only be able to access forms that are in the projects they have been given access to. This can be used to prevent enumerators and analysts access forms that they are not working on. However users with the admin security group can always give themselves access to any project

Project Management

Forms are stored in projects. Users can be given access to a project which gives them access to the forms in that project

- f. Use projects to:
 - i. Simplify the work environment
 - i. Enumerators would generally only have access to the project they are working on so they are not distracted by other surveys

- ii. Analysts may have access to several projects but only in the area they are working on
 - iii. Keep the number of forms per project small
 - ii. Security
 - i. Users who do not have the “admin” role will not be able to look at forms in projects they have not been assigned
- g. Admin can Create, Delete and Rename projects on the User management page.
- h. To get to the user management page, Select the “Users” menu from inside the “admin” module

	Id	Name	Changed by
<input type="checkbox"/>	188	UNOPS_Post_Training_Test3	rajeevherd
<input type="checkbox"/>	165	UNOPS_v2	santosh
<input type="checkbox"/>	311	UNOPS-WT	rajeevherd
<input type="checkbox"/>	548	urban-health	rajeevherd

- i. To create a project click on the “Add Project” button
- j. To rename a project click on the button with the project's name
- k. To delete a project click on the check box next to the projects you want to delete then press the “Delete” button under the list of projects
- l. When admin create a project admin users are automatically added as a user of that project

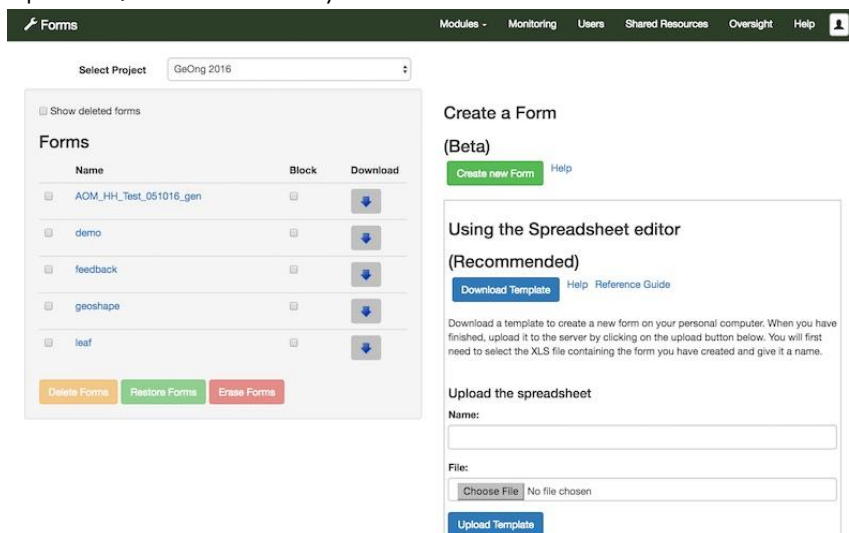
Survey Form Management

Form Management page provides following features.

- Upload / Create and Delete forms
- select a form for editing
- recover deleted forms and permanently erase them
- download the form as an XLS Form, PDF or Codebook
- Block submission of data to a form

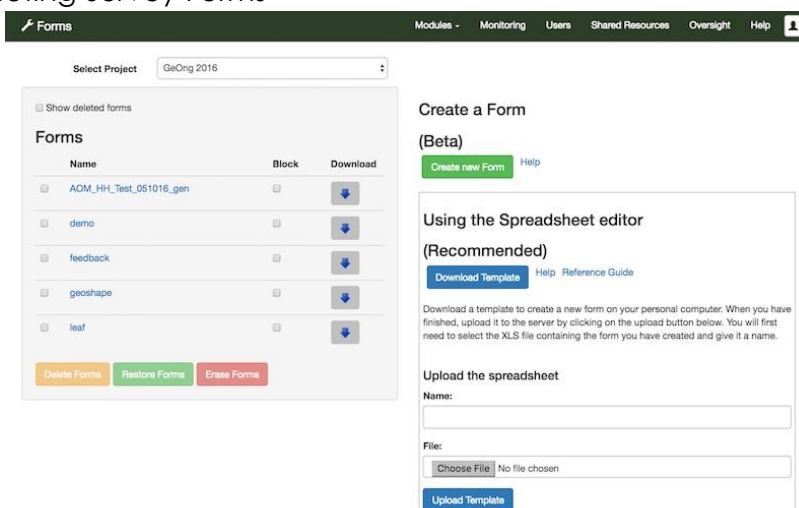
User must need to have the “Admin” or “Analyst” security group in order to use this page. Form Management page is the default page when user logged in in sg.smap.com.au server. This page can get to it by select Modules and the then Admin.

a. Upload / Create survey Forms.



- i. We can create a new Form using the online editor by clicking the green button. However we generally use excel spreadsheet to digitize paper based tool and it is more efficient and convenient.
- ii. Digital spreadsheet an ODK XLS Form Excel spreadsheet can be uploaded.
 - a. Click “Choose File” button and select excel file.
 - b. Type title of survey form under the field “Name”.
 - c. And click “Upload Template” button to upload form.

b. Deleting Survey Forms



- i. The list of current forms is shown on the left of the page.
- ii. To delete one or more forms first select the checkbox at the left of the form name.
- iii. Then press the delete button.

The form will disappear from the list and we will no longer be able to access the results or submit new results. However the deleted form can be restored so that you regain access to its data.

c. Restore Deleted Survey Forms

Name	Block	Download
<input type="checkbox"/> AOM_HH_Test_051016_gen	<input type="checkbox"/>	
<input type="checkbox"/> demo	<input type="checkbox"/>	
<input type="checkbox"/> feedback	<input type="checkbox"/>	
<input type="checkbox"/> Feedback (2016-10-15 21:13:13)	<input type="checkbox"/>	
<input type="checkbox"/> geoshape	<input type="checkbox"/>	
<input type="checkbox"/> leaf	<input type="checkbox"/>	

- i. To show deleted forms in the list select the “Show Deleted” checkbox.
- ii. Deleted forms will then be shown in red. They will have the date and time that they were deleted added to their name.
- iii. To restore a form, select the checkbox next to its name and press the green restore forms button. You will then be able to view submitted data and submit new data to it.

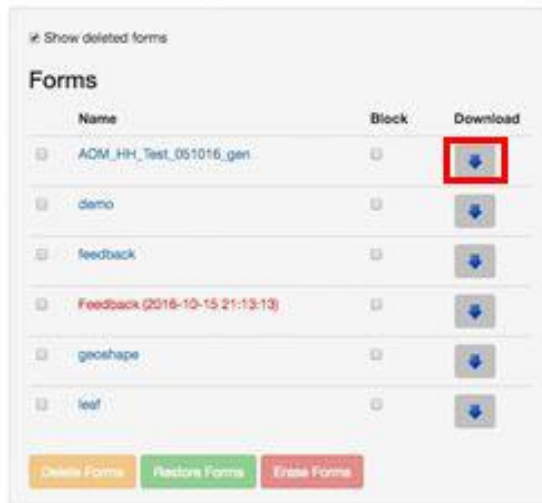
d. Erase Deleted Survey Forms.

- i. You can only erase deleted forms hence you will need to click the checkbox to show deleted survey form first.
- ii. To erase a form, select the checkbox next to its name and press the red erase button. The data will then be deleted from the system and you will not be able to recover it without using a backup.
- iii. If there is any data uploaded to the form the system will show you the number of records and ask you to confirm the erase.

e. Block uploading data to Survey Forms

Name	Block	Download
<input type="checkbox"/> AOM_HH_Test_051016_gen	<input type="checkbox"/>	
<input type="checkbox"/> demo	<input type="checkbox"/>	
<input type="checkbox"/> feedback	<input type="checkbox"/>	
<input type="checkbox"/> Feedback (2016-10-15 21:13:13)	<input type="checkbox"/>	
<input type="checkbox"/> geoshape	<input type="checkbox"/>	
<input type="checkbox"/> leaf	<input type="checkbox"/>	

- i. Click the checkbox next to the form name in the “blocked” column
 - ii. We may want to block uploading of new data if you have started your analysis and begun cleaning the data.
 - iii. If an enumerator tries to submit results from their phone to a blocked form they will get a “Not Found” message.
- f. Download uploaded survey forms in different formats.



- i. Select the Download button next to the file name.
- ii. A dialog box will ask you the type of file you want.

File Type

- XLS Form (Version: 1)
- Original XLS file (Version: 1)
- XML Form
- Codebook
- PDF Template

- i. The first XLS Form is the latest version of the survey after any changes have been made in the online editor.
- ii. If you uploaded the form from an XLS Form then you can also download that original file.
- iii. XML Form is a format used by ODK and may be useful if you are going to upload the form into another ODK system.
- iv. Codebook will contain the choice names and labels that you can reference when doing your analysis.
- v. PDF template will be easy to read version of the survey that you could print out and use as a paper form.

Monitor uploaded survey form

- a. Select "Admin" then the "Monitor" menu option
- b. The default monitor view is to show the forms that have been downloaded to fieldTask android based application.

Menu ☰

Monitoring					Rajeev Dhungel @HERD
Source <input checked="" type="radio"/> Forms <input type="radio"/> Submitted <input type="radio"/> Forwarded	User Name	User Ident	Device Identifier	Survey Name	Survey Version
	Freddy	freddy1	undefined	Not Downloaded	
	Rajeev Dhungel	rajeev	9774d56d682e549c	food	1

- i. This view shows the user that has downloaded the form, their device, the form name and the form version
 - ii. It's the default view when the monitor page is opened, Set the source to "Forms" if another view is being shown
 - iii. In this example
 - iv. User freddy1 had not downloaded any forms for this project even though they have enumerator rights in the selected project
 - v. User rajeev has downloaded the form "food", version 1 to one device
 - vi. If a user has downloaded forms to more than 1 device then all of those device will be shown
- c. This view shows the total number of submitted results for each survey in the project

Monitoring					Rajeev Dhungel @HERD
Source <input type="radio"/> Forms <input checked="" type="radio"/> Submitted <input type="radio"/> Forwarded	Survey	Success	Errors	Duplicates	Not Loaded
	CFW PDM Field Test Sept26	51	0	0	0
	CFW PDM Sept2014	217	0	0	0
	Final CFW-PDM Sep2014	384	0	3	0

Show
 Totals
 Instances (last 200)

- i. Set the source to "Submitted".
- ii. Set the "show" to "Totals".
- iii. In this example
 - a. There are 3 surveys with submitted results in the project.
 - b. There are no failed uploads
 - c. The final CFW-PDM survey has 384 sucessfully loaded submissions. There were also 3 duplicates detected which were discarded.
- iv. Note duplicates can usually be ignored. They are caused by an error during submission of result when the server has received results correctly but the phone does not get the message confirming that. Hence the results are sent again. If there are a

lot of duplicates it may mean that there is a configuration problem with your server and timeouts need to be changed. (IT support issue)

d. This view shows the individual submissions

Monitoring											Rajeev Dhungel @HERD
Source	Id	Upload Id	Upload Date	User	Phone Identifier	File with raw results	Survey	Ident	Location	Status	Failure Reason
Forms Submitted Forwarded	34766	34501	2014-10-03 10:53:20.563+00	haiyan	864718012450864	7096b7eb-59e7-4e03-aa44-b1ad45213f1c.xml	Final CFW-PDM Sep2014	s51_1296	none	success	
Show <input type="radio"/> Totals <input checked="" type="radio"/> Instances (last 200)	34765	34500	2014-10-03 10:53:20.028+00	haiyan	864718012450864	b4ec1dd2-9696-4200-bdfc-926f6df82e6f.xml	Final CFW-PDM Sep2014	s51_1296	none	success	

- i. This view shows the individual submissions
- ii. Set the source to "Submitted"
- iii. Set the "show" to "Instances"
- iv. Note only 200 submissions are shown. If you filter out the successful submissions and only show errors then the last 200 errors will be shown

Analysis of uploaded data.

This section provides the detail overview of submitted data collected using mobile survey.

- e. Online view data in different representable format.
 - i. Table
 - ii. Graph
 - iii. Map
 - iv. Media
- f. Retrieval / exporting method to analyze data offline.

Please refer attached document for this purpose.

a. Analysis

b. Advanced Analysis



Advanced Analysis.pptx

Relevant documents

Other useful documents can be retrieval from following links.

- a. https://drive.google.com/drive/folders/0B5_SmpWlQYxvNExwSFV5X051Yjg
- b. <http://smap.com.au>
- c. <https://www.youtube.com/user/ianaf4you>

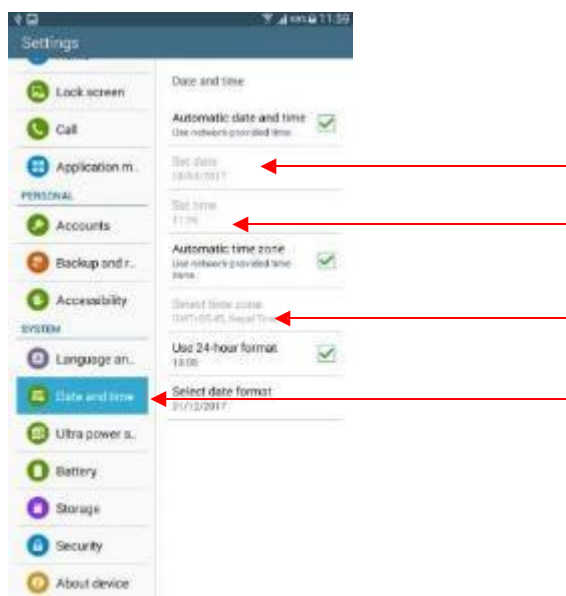
How To: Setup tablets

Settings

Date and time.



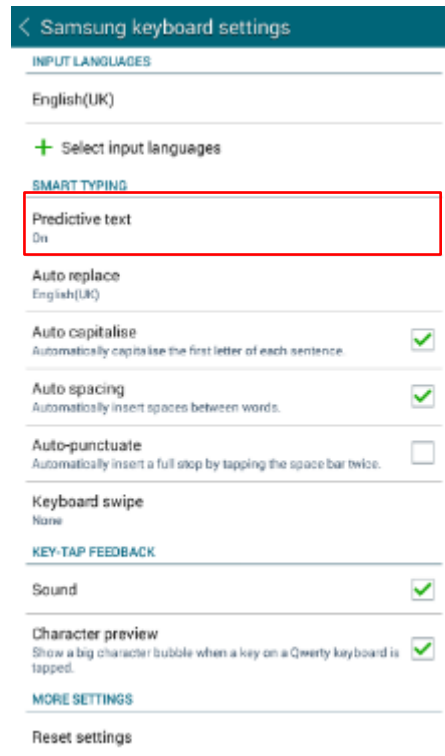
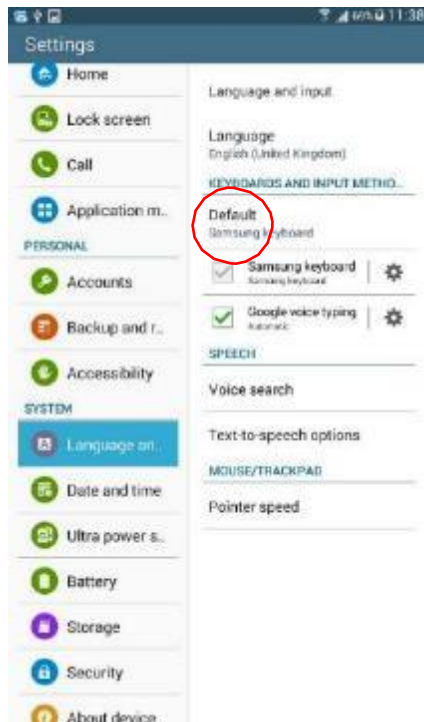
a. Tap on date and time



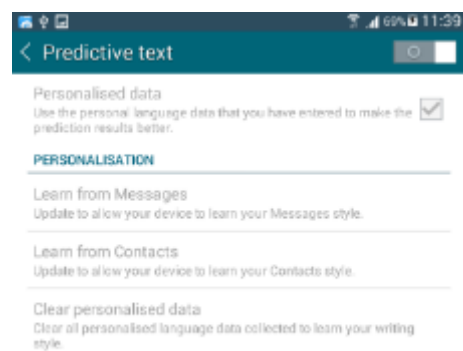
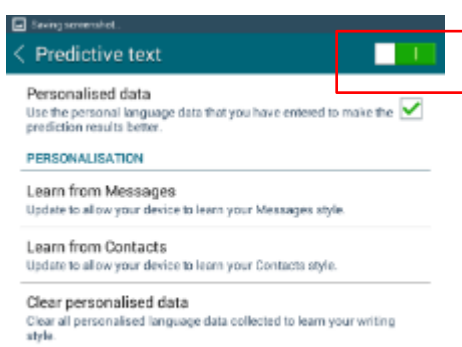
b. Change each field if seems incorrect value.

Set prediction off.

- a. Select Samsung keyboard setting
- b. Tap on prediction



- c. Tap on the off / on button to **off** predictive text

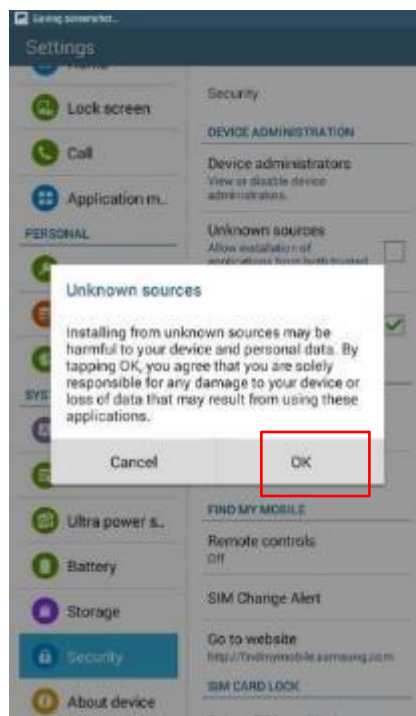


Allow third party applications.


- a. Go to setting
- b. Scroll Down to security
- c. Check on unknown sources



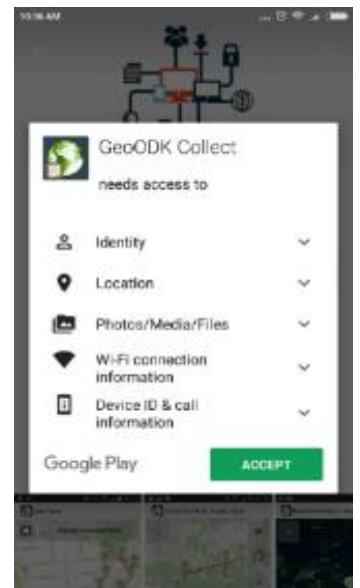
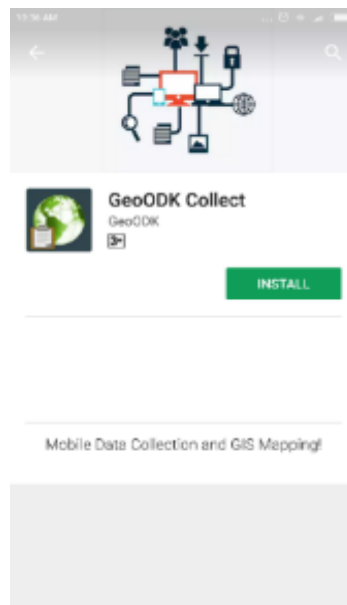
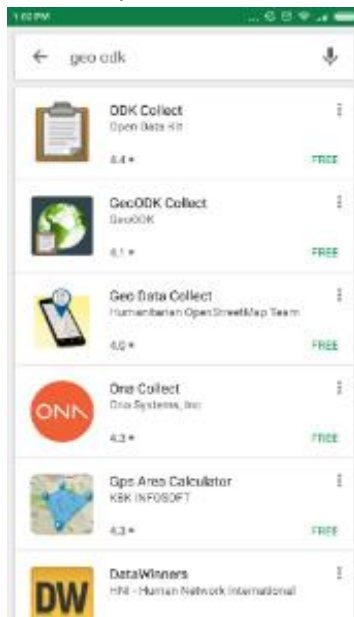
- d. Prompts warning
- e. Tap on ok



Install applications

- Go to play store 
- Search for each of the following apps: wifi manager, App Lock, MAPS.ME, Vespucci OSM Editor, GeoODK Collect, Maps & GPS Navigation (OSMAnd), Nova Launcher, CamScanner-phone PDF Creator
- Tap on install
- Accept the app

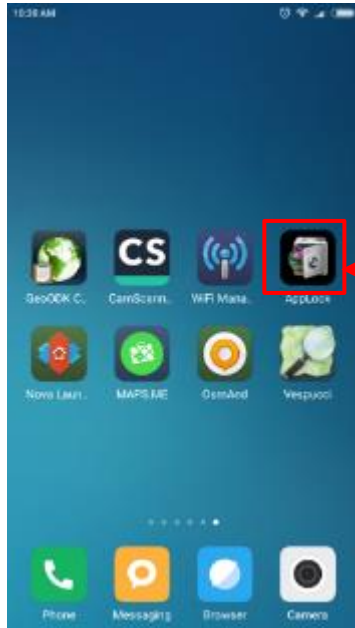
For example:



Configure applications

App Lock

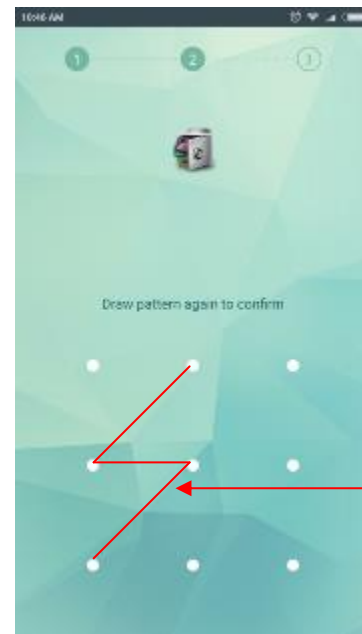
1. Tap the app lock app
2. It will prompt you with initial pattern password
3. confirm the pattern password



1. Tap on app lock

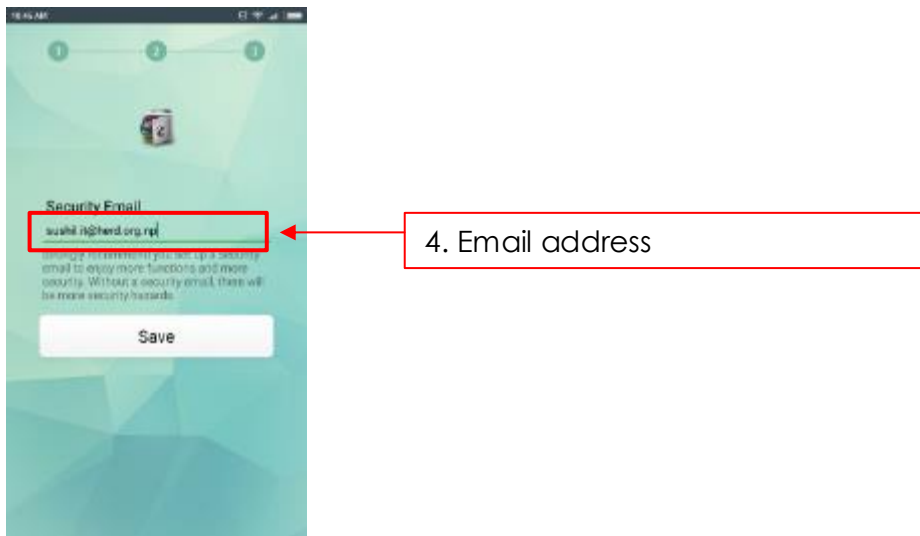


2

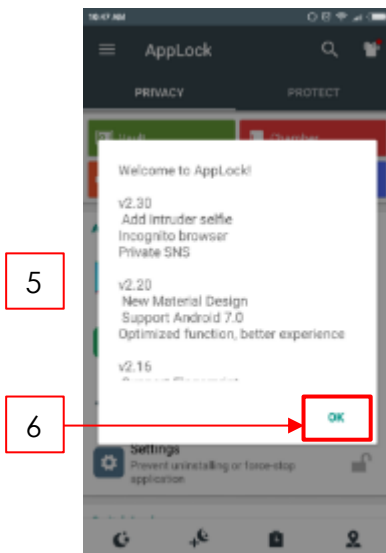


3

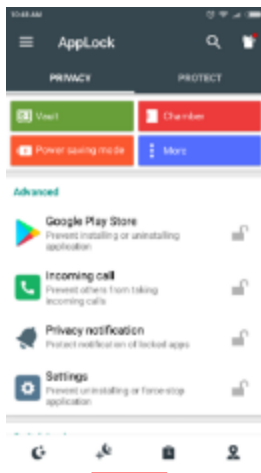
4. After confirming the pattern password, it will prompt with password recovery method and provide field to put email address. Enter email address (herdint.camscan@gmail.com)



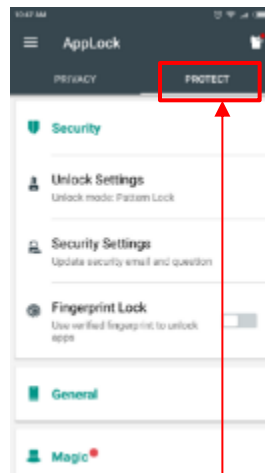
5. It will prompt for welcome
6. Tap on ok



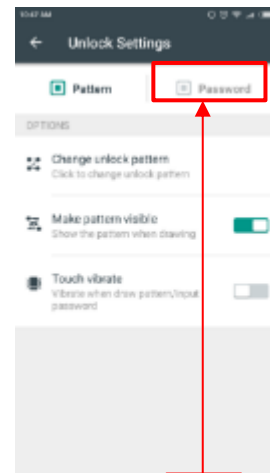
7. It will show you home screen of app lock
8. Tap on protect
9. Tap on password
10. Type the password (In our case we use 9632)
11. Retype the password (9632)
12. Provide hint as shown in picture.



7



8



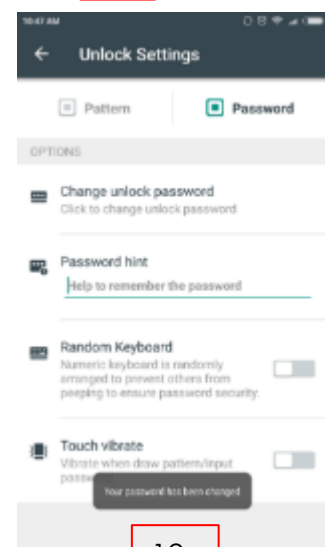
9



10

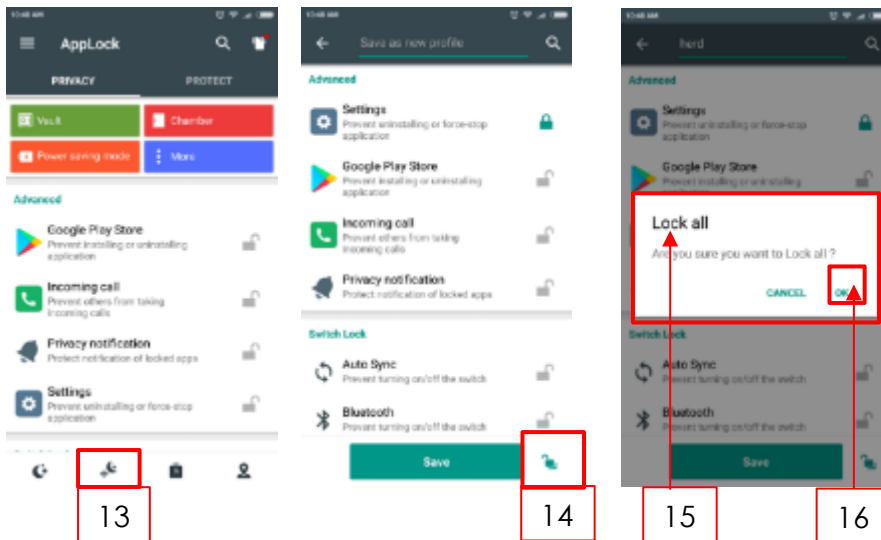


11



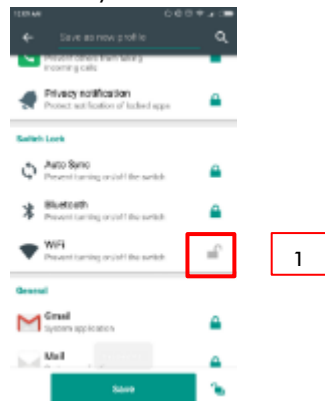
12

13. Tap on new profile tab
14. Tap on lock all icon
15. Prompt you with "Are you sure you want to lock all"
16. Tap on ok
17. Tap the desire app you want to give unlock permissions by tapping on unlock icon



18. Unlock permissions should given to following apps

- ✓ Wi-Fi
- ✓ WiFi Manager
- ✓ CamScanner-phone PDF Creator
- ✓ Maps & GPS Navigator –OsmAnd
- ✓ MAPS.me
- ✓ GeoODK Collect
- ✓ Camera
- ✓ Location
- ✓ Gallery

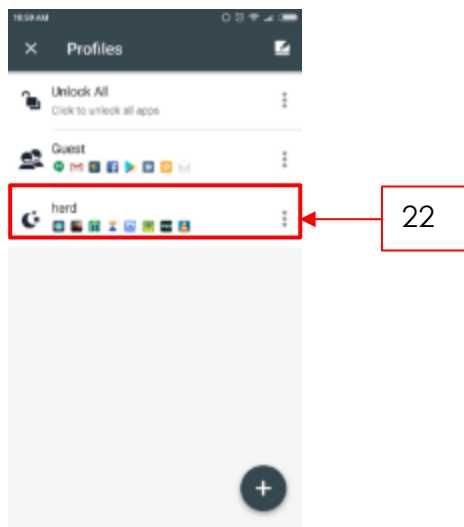
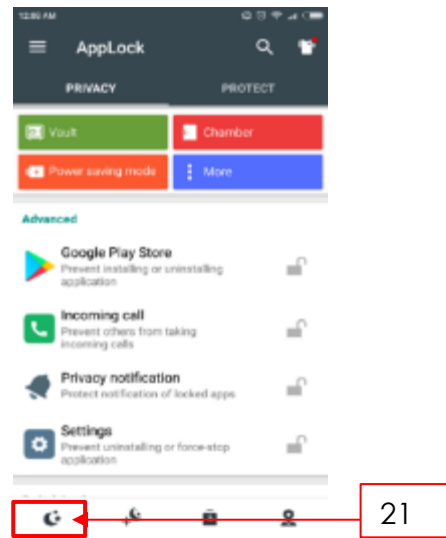
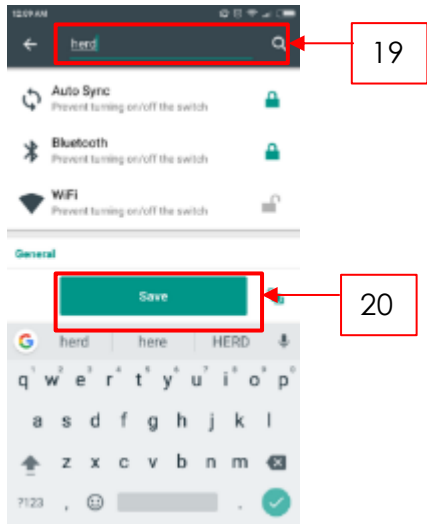


19. Name the Profile (in our case we use herd)

20. Tap on Save

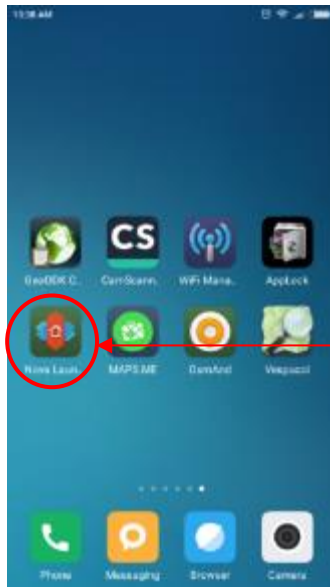
21. Tap on profile tab

22. Select the profile you created



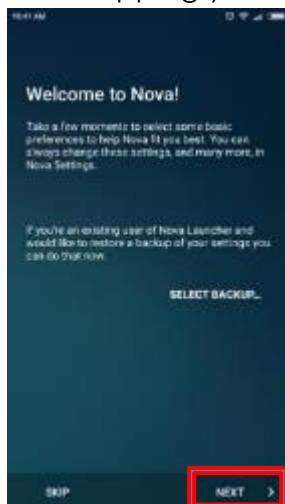
Nova Launcher

1. Tap Nova Launcher application.



1. Nova Launcher

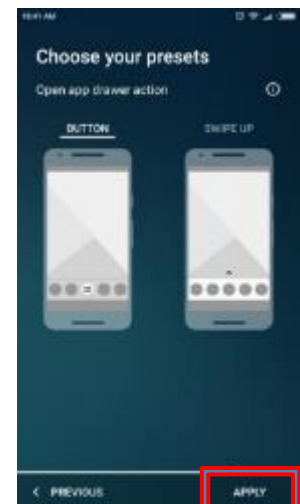
After Tapping you will See this Screen



2



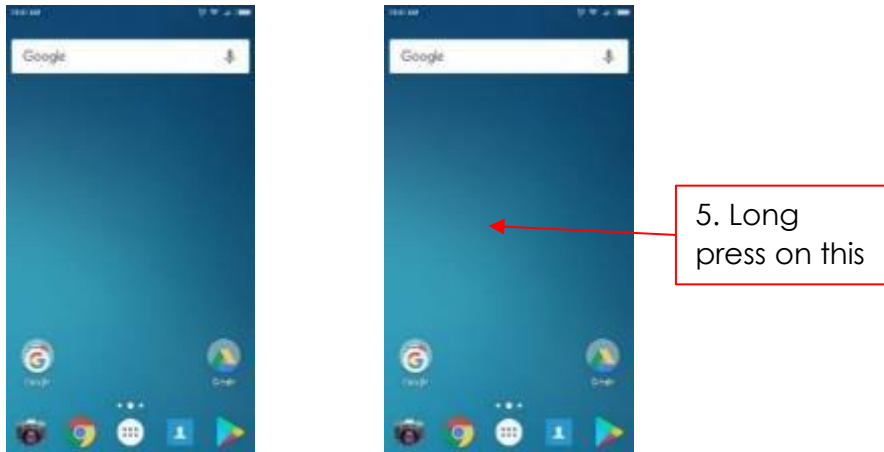
3



4

2. Tap on next
3. Tap on next
4. Tap on Apply

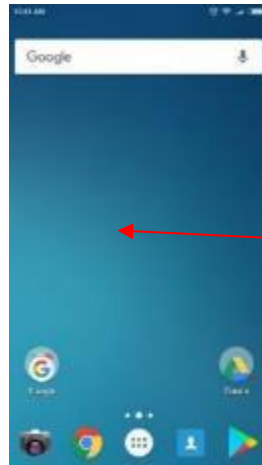
After tapping Apply you will see these screen



5. Long press on the Screen
6. After pressing long on the screen you will see this screen
7. Scroll to left
8. This is blank screen
9. Press on + on the Screen
10. You will prompt Warning
11. Tap on ADD BLANK



12. Long press on the Screen
13. After pressing long on the screen you will see this screen
14. Long press on the Screen
15. Move to upper remove icon
16. You will remain this screen



12. Long press on this screen



13



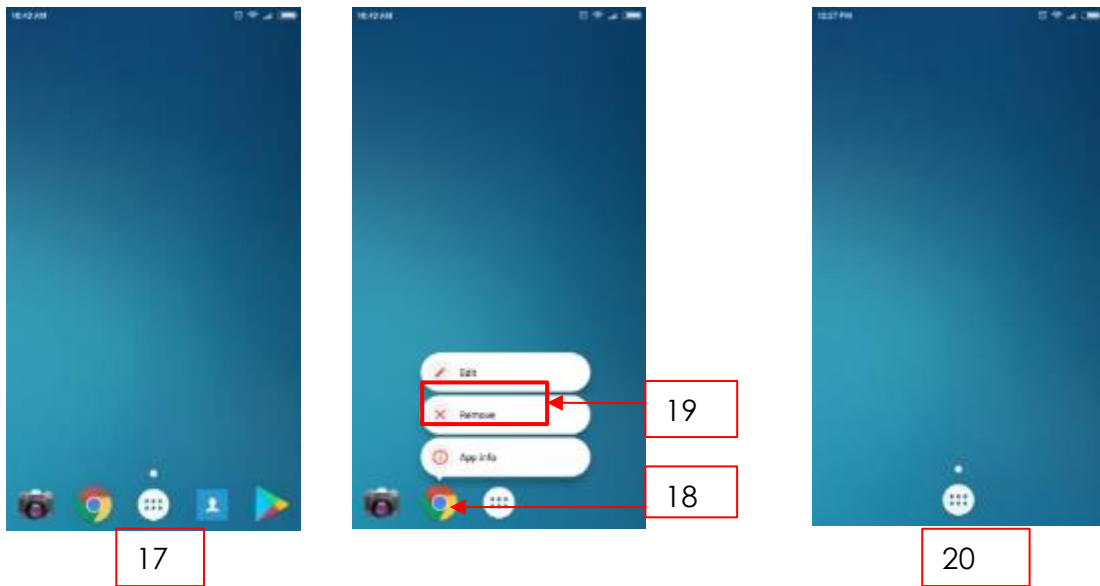
15. Remove

14. Long press on

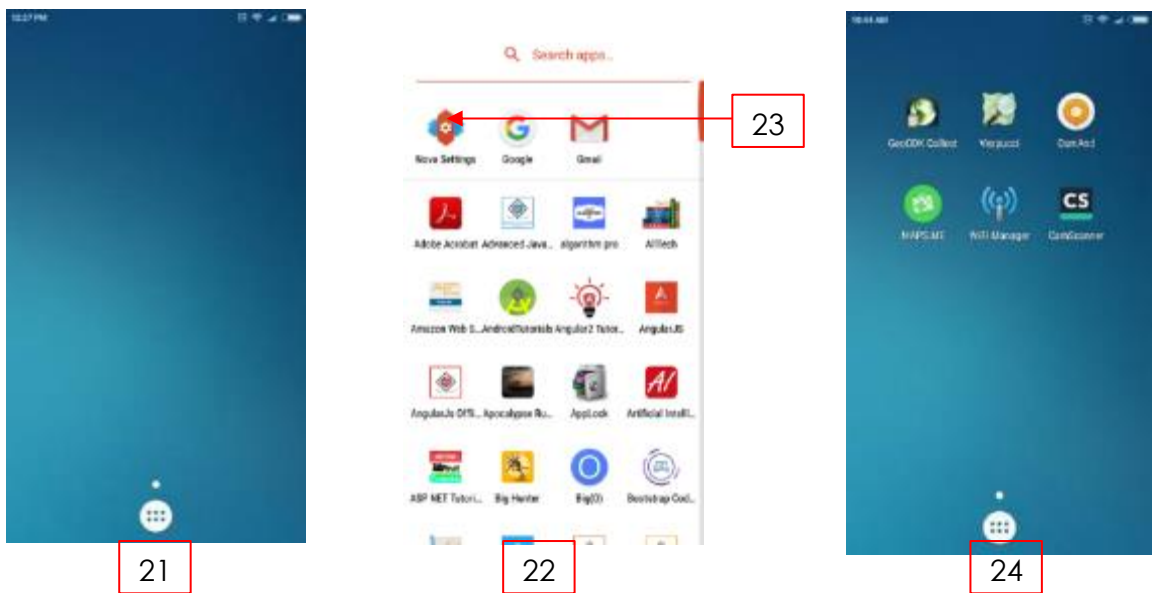


16

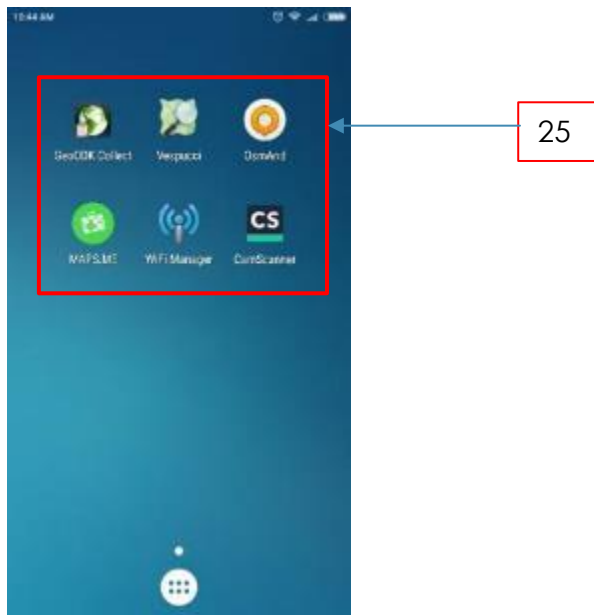
- 17. You will be with this screen
- 18. Long press to the one icon at a time
- 19. Tap on remove
- 20. You will be with this screen



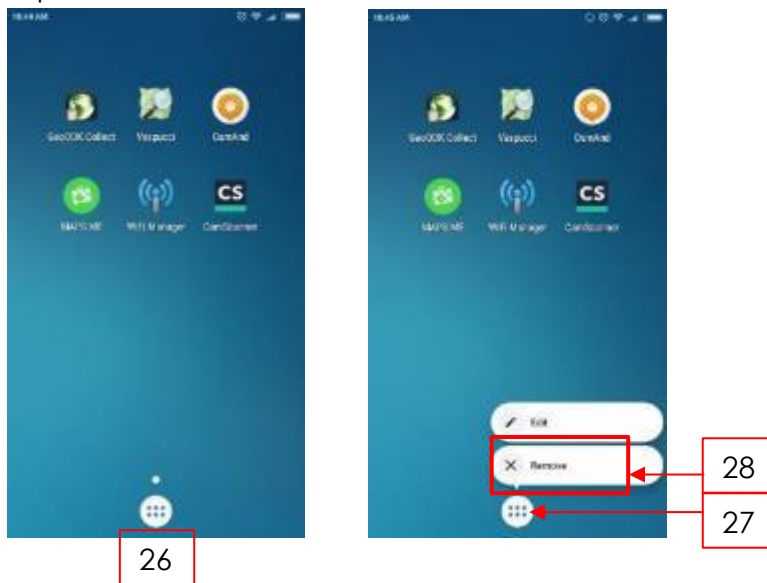
21. Tap on home menu
22. You will find this screen
23. Long press the app icon
24. It will make short cuts of app icon at the home screen



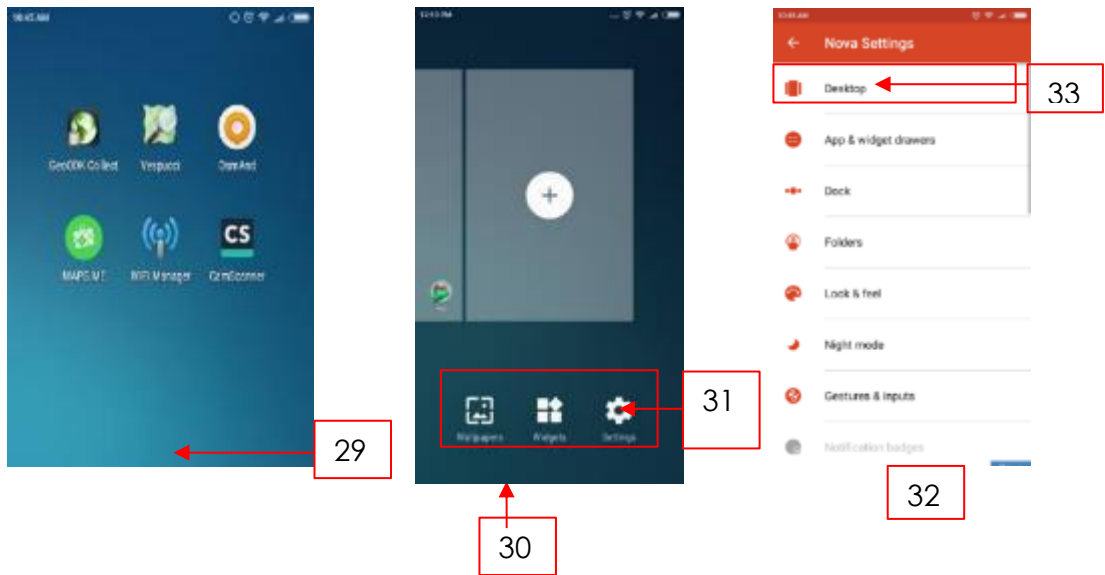
25. Make Shortcuts of application as mention below.
 - ✓ WiFi Manager
 - ✓ CamScanner-phone PDF Creator
 - ✓ Maps & GPS Navigator –OsmAnd
 - ✓ MAPS.me
 - ✓ GeoODK Collect



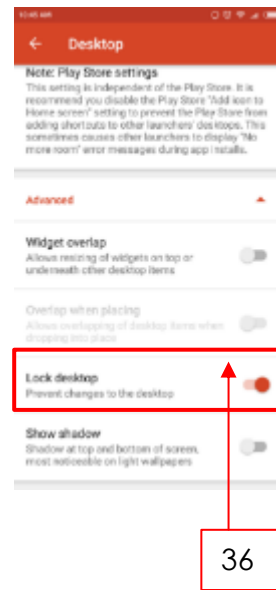
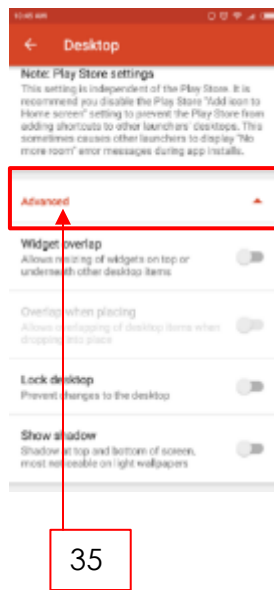
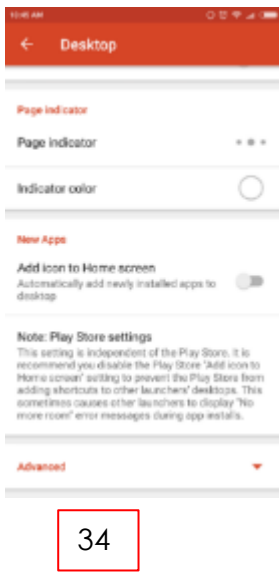
- 26. You will have left with this screen
- 27. Long press to home button
- 28. tap on remove



- 29. long press on the screen
- 30. Some menu will appear on the screen at the button
- 31. Tap on Setting
- 32. Nova Setting will appear



- 33. Tap on Desktop
- 34. Scroll down and find Advanced
- 35. Tap on Advanced
- 36. Tap on and enable Lock desktop



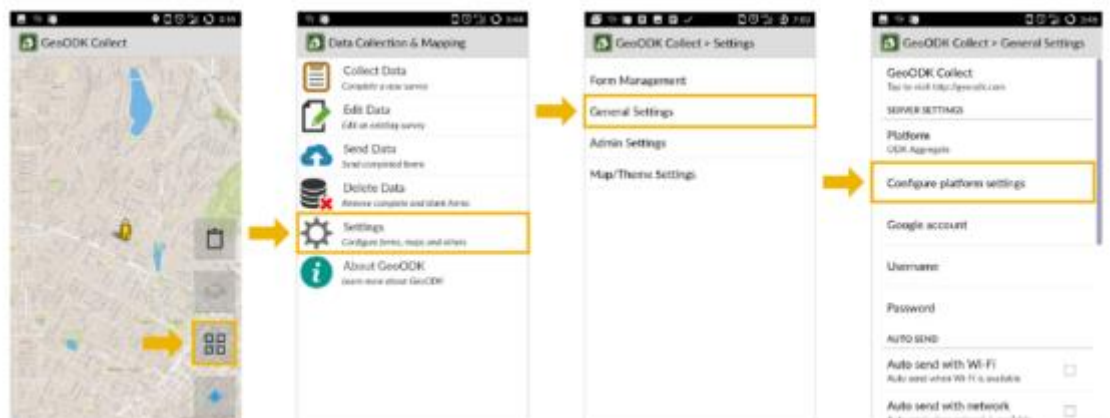
Geo ODK

Platform setup

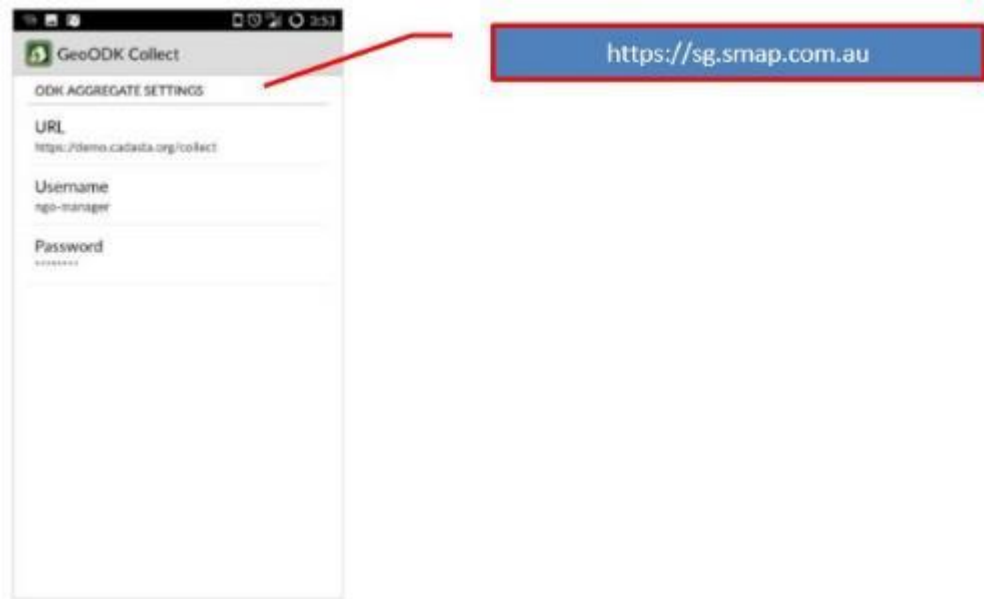
- a. Tab Geo ODK application.



- b. From the map screen, hit the button with the four squares on the right. Then select **Settings** from the Main Menu, then **General Settings**, then **Configure Platform Settings**.

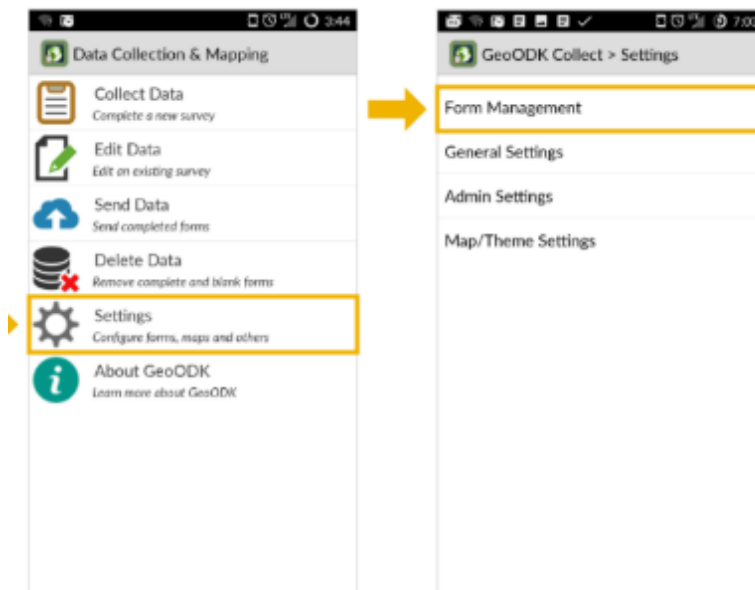


- c. On this screen, enter the Platform URL along with your username and password. Now you have a GeoODK account that is synced with the platform.



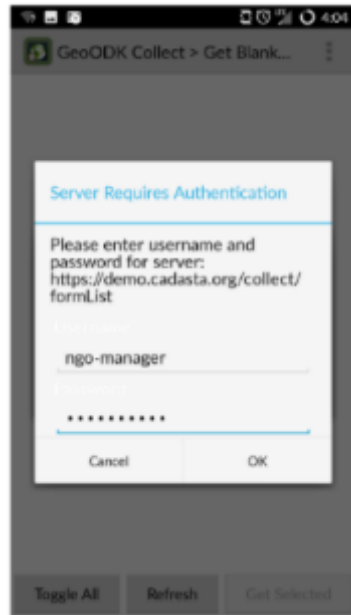
Loading your XLS form

- a. Once you are connected with the platform, you need to download the form you are using for data collection.
- b. From the main menu, select **Setting** and then **Form management**.



- c. At this stage you may be asked to enter your platform username and password. Click Ok button.

Note:- Please don't modify username and password



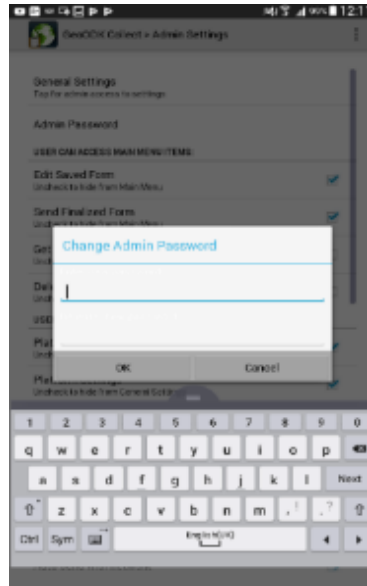
- d. In the page that follows, you'll see a list of XLS Forms that have been loaded for your organization's projects. Checkmark the form you'd like to download and tap **Get Selected**.



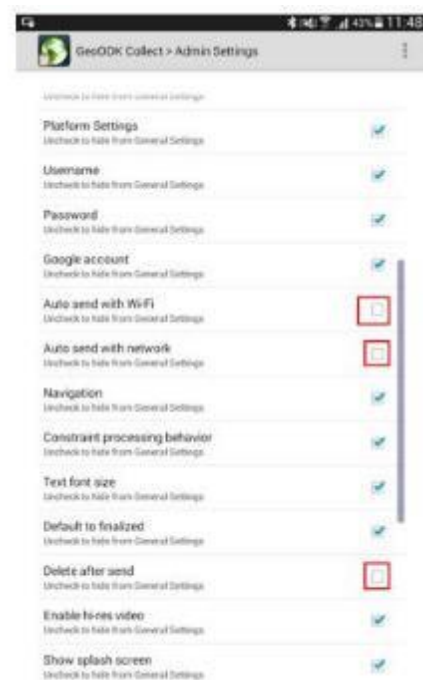
- e. Now, GeoODK is configured to collect data using your data collection form.

Restriction

- a. Enumerators should be restricted using certain functions/features. First, you need to set admin password. From the map screen, hit the button with the four squares on the right. Then select **Settings** from the Main Menu, then **Admin Settings**, then **Admin Password**.

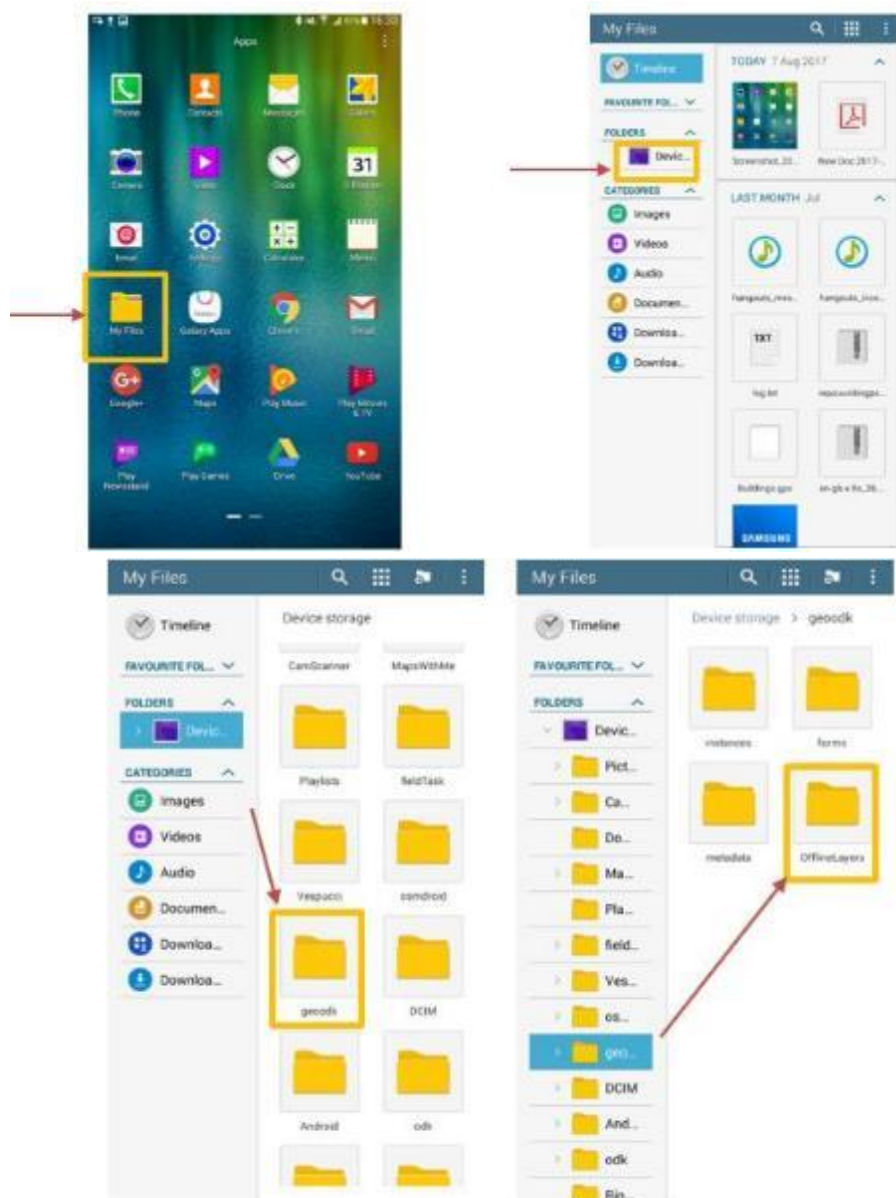


- b. Restrict some features to enumerators. Uncheck **Get blank forms**, **Delete saved forms**, **Auto send with Wi-Fi**, **Auto send with networks** and **Delete after send**. These functions will not show in general settings for enumerators.



Offline Layer

- Create gpx file. Refer to gpx file creation manual.
- Tab My Files.
- Click on device. Open GeoODK and OfflineLayers folder



- d. Create new folder named as per convient (SUE-PSU) by clicking icon as shown in fig below

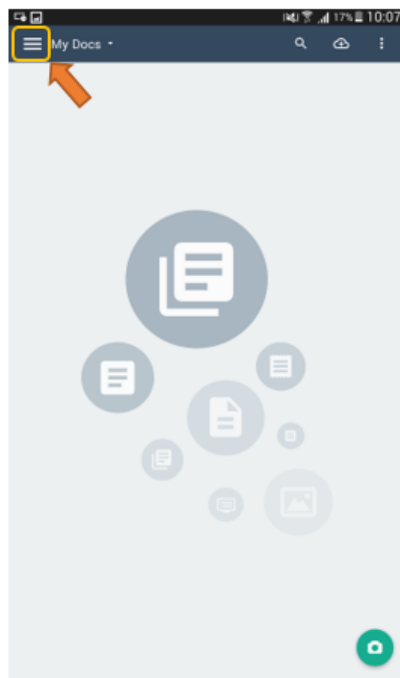


CamScanner

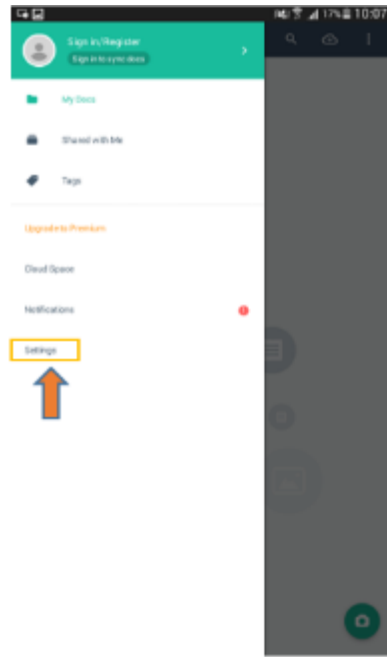
- a. First Select the Cam Scanner app in tablet
- b. Tap to Open Cam Scanner app in tablet



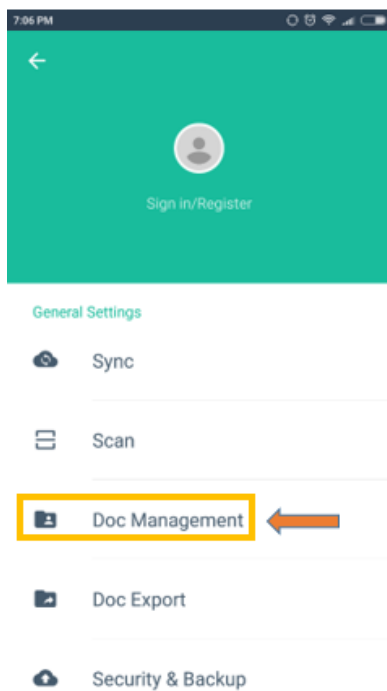
- c. You will see this Screen and tap the left top corner



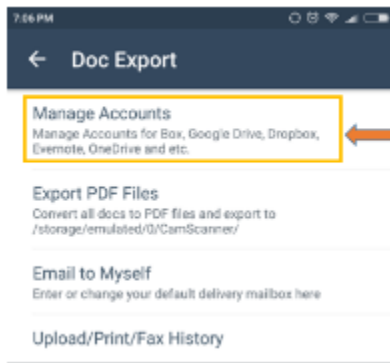
d. Screen appears as below and click settings.



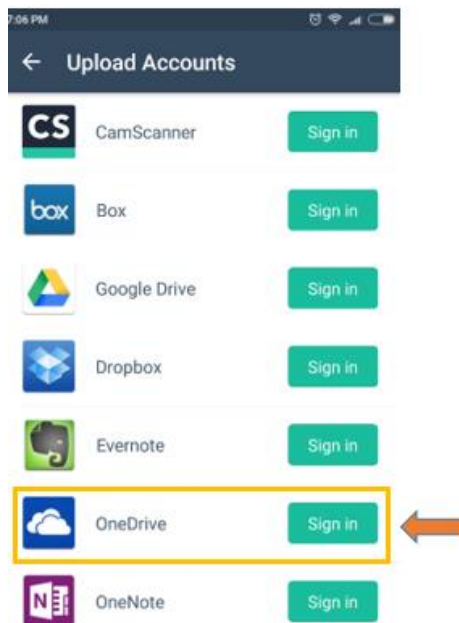
e. Screen appears as below and click on Doc Export



f. Tap on Manage Accounts



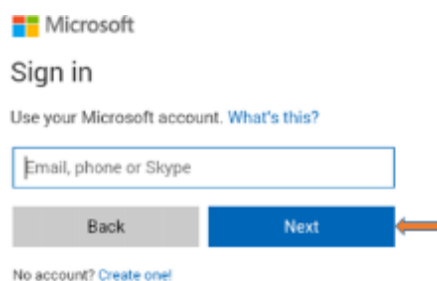
g. You will See This Screen and select OneDrive and sign In

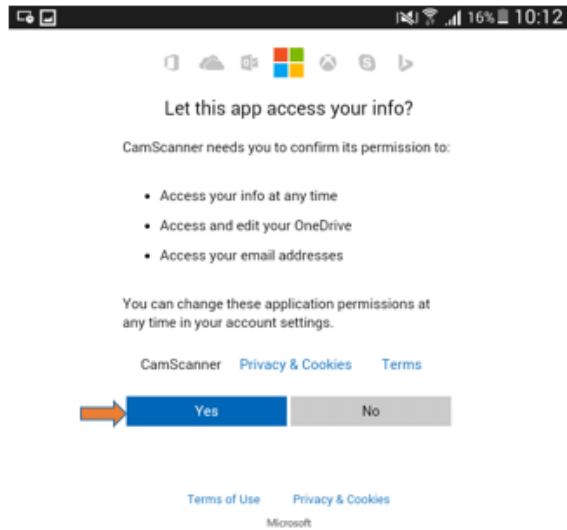


h. Sign in with OneDrive in our case we use following credential.

Username: herdint.camscan@gmail.com

Password: pas\$w0rD123





- i. Select Yes.

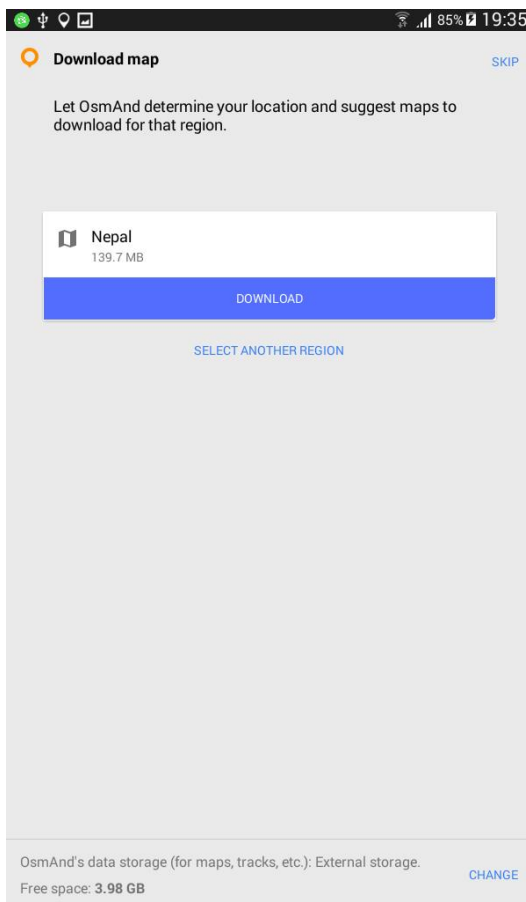
MAPS.ME

Step 1: In Google Play, search for “MAPS.ME” and install. Accept that the app will access your location. Select “Open” after the app installs. Then select “Accept and continue”.

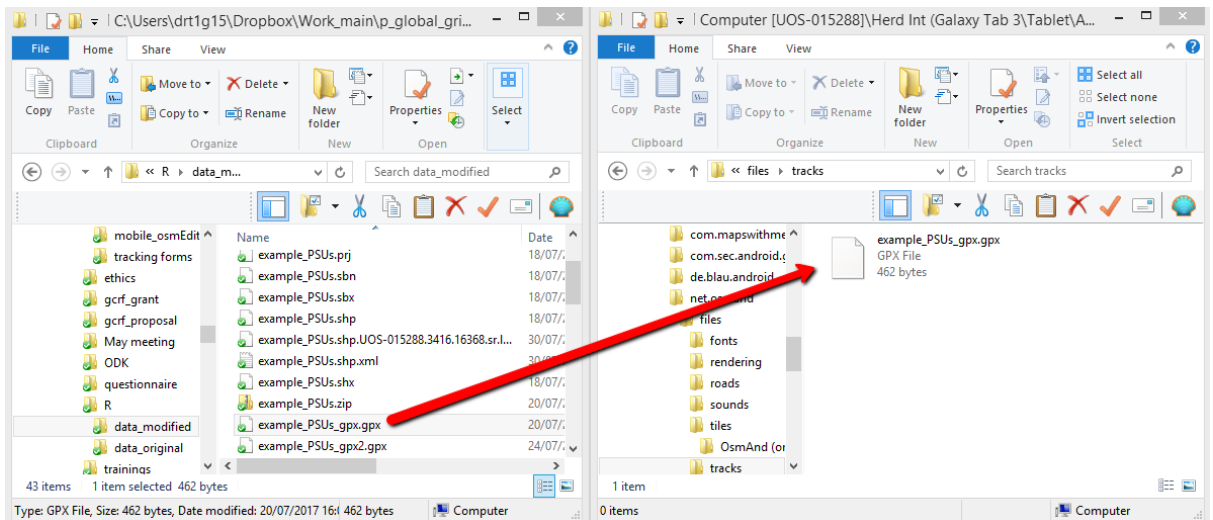
OSMAnd


Step 1: In Google Play, search for “OSMAnd” and install “Maps & GPS Navigation – OSMAnd”. Accept that the app will access your location and photos/media/files.

Step 2: OSMAnd will detect that you are located in Nepal (or another country) and offer to download maps for your country. Accept to download relevant national and international maps. This will take a few minutes. Click “display on map” when done.

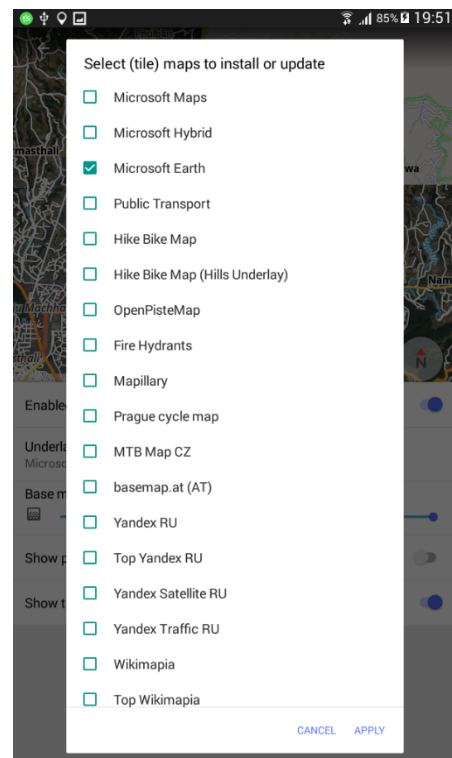


Step 3: To complete this step, you will need to have previously converted the PSU boundaries to GPX tracks in ArcGIS or QGIS. See ref_gpx.pdf. Save the GPX tracks to the “Tablet\Android\data\net.osmand\files\tracks” folder on your device.

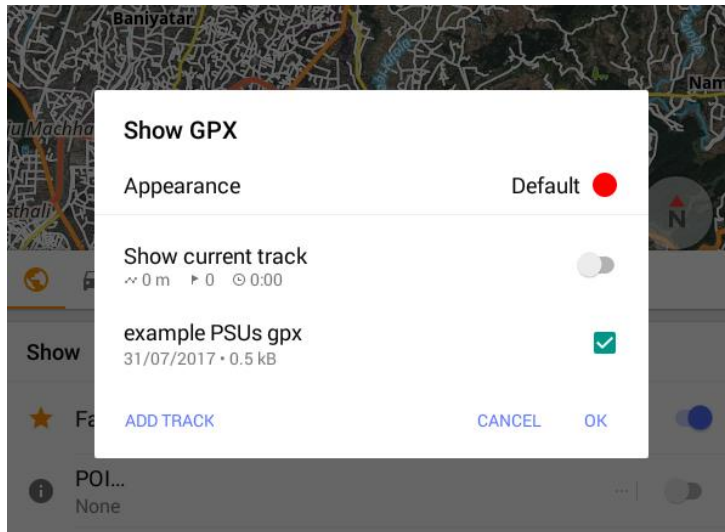


Step 4: Modify the following settings by selecting the 3-bar menu  in the bottom left corner, then “configure map”.

- To enable background layer, select “Map source...” and enable access to online satellite imagery by touching the word “Disabled” until it says “Enabled”. Return to the main screen when you are done.
- To add a background layer, select the globe in the upper left corner, and select “underlay map”. Turn the toggle to “enabled”, select “install more”, choose “Microsoft earth”, and then “apply”. Return to the main screen when you are done.



- To display PSU boundaries: Select the 3-bar menu in the bottom left corner, “configure map”, “GPX tracks”, and “add tracks”. Select the appropriate tracks file. Return to the main screen when you are done.

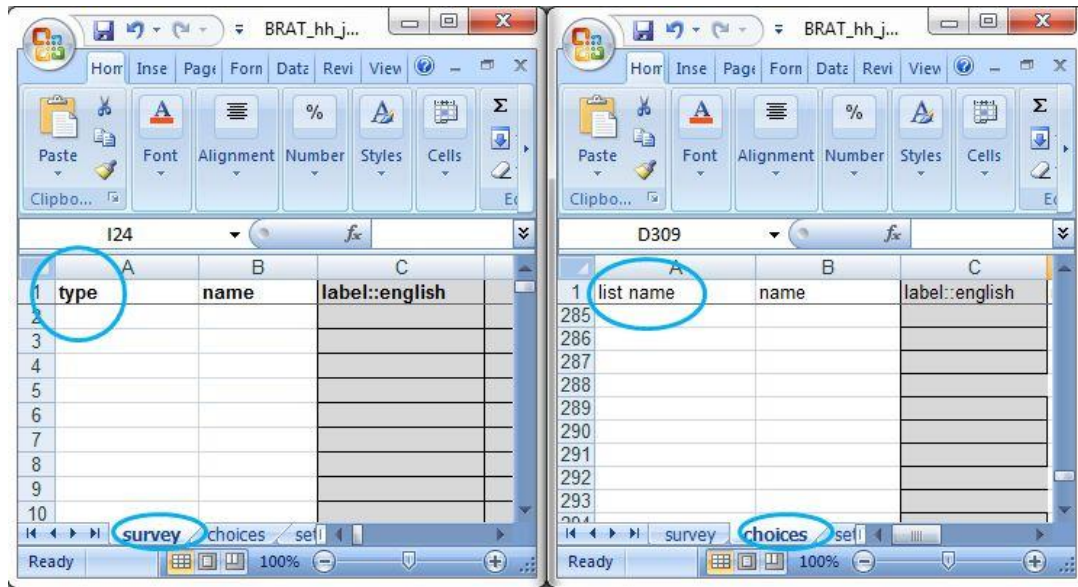


- If you want your enumerators to be able to add points of interest to the map from the field, select the 3-bar menu in the bottom left corner, “plugins”, “OSM editing”, and change from “disabled” to “enabled”. Enumerators will have to enter their OpenStreetMap username and password to submit edits through this app.

How To: Setup GeoODK form in Excel

Excel document preparation.

For the preparation of digital tool which is being created using excel document. The spread sheet has two sheets.



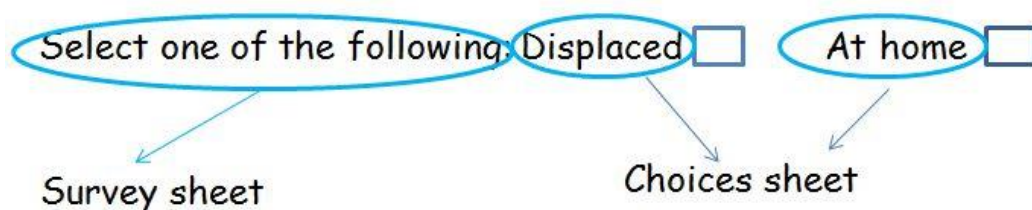
Survey

Choices

The "survey" sheet is where we put the questions themselves. It allows us to customize set the type of questions. For example: "select_one" or "select_multiple" or "integer". As well as constraints, skip rules and other question attributes.

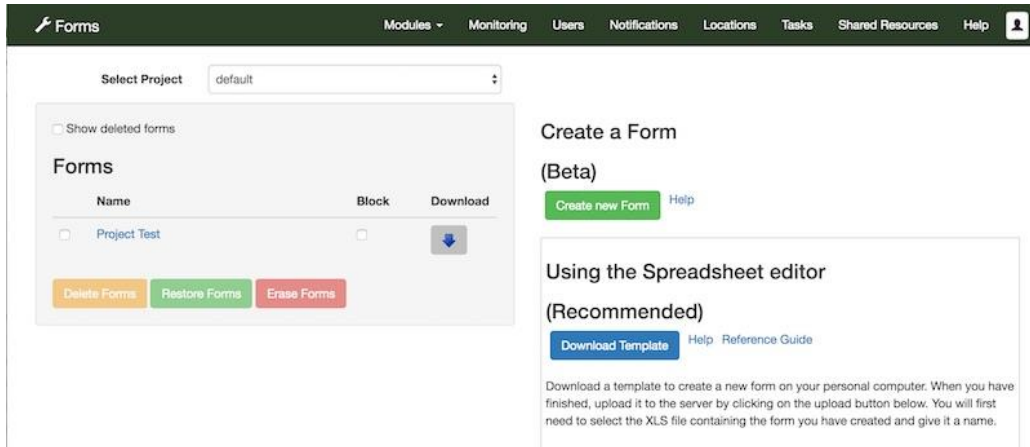
The "choices" sheet, as indicated by its name, is where all the different response options are listed. So in a "select_one" question the list out of which a single answer is selected will be found on the choices sheet.

For example in a question that looks like the following in a paper document will be partly specified in the survey sheet and partly in the choices sheet.



Download survey template Excel file from SMAP server.

- Browse <http://sg.smap.com.au>
- Login with username and password.



- Click download template button to download standard survey template designed in excel file.

Example to create survey form

a. Text field

Paper based tool where Name of the surveyor question is asked.

[Response]:	
1. Assessing agency information	
Name of the surveyor	Date of assessment
	Contact information of surveyor

When digitizing this question it can be represent as shown in fig below.

type	name	label
text	surveyor_name	Name of the surveyor

Type: - This field is the attribute of question which might be text, integer, select one, select multiple, etc.

Name:- This field is unique name for the question. It doesn't allow spaces and allowed English letters and numbers and should start with a letter.

Label:- The question shown on the screen, can be in any language.

On the phone text, integer and characters can be entered in response to this question

b. Integer field

Paper based tool where Total Number of People question is asked.

3. Demographics		
3.1	What is the total number of people <i>currently</i> with you here in your household? (including yourself)	
3.2	Compared to before the emergency, do you have <input type="checkbox"/> more <input type="checkbox"/> the same <input type="checkbox"/> fewer people living in your household now?	

When digitizing this question it can be represent as shown in fig below.

type	name	label
integer	number_in_hh	What is the total number of people currently with you here in your household (including yourself)?

On the phone only numbers can be entered in response to this question

More training

HERD International offers a one week detailed training on developing tablet based survey forms. Please contact Rajeev at HERD International for more information.

Reference Documents

Other useful documents can be retrieval from following links.

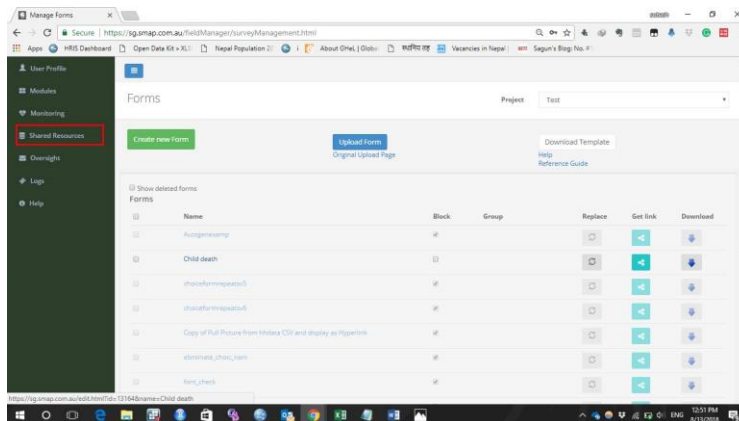
- d. https://drive.google.com/drive/folders/0B5_SmpWlQYxvNExwSFV5X051Yjg
- e. <http://smap.com.au>
- f. <https://www.youtube.com/user/ianaf4you>

How To: Pull listing data into questionnaire form

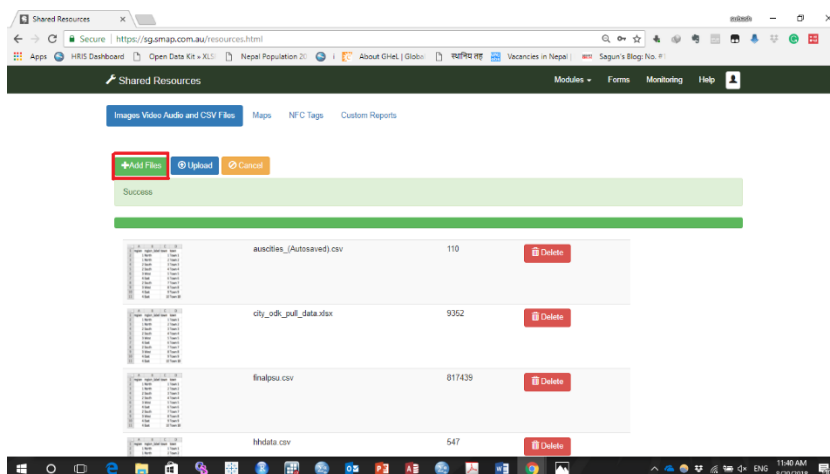
- 1) Browse URL **sg.smap.com.au**
- 2) It pop up for login. Use credential to use **“log in”**.
- 3) Click on **“Admin”** as shown in the picture below.



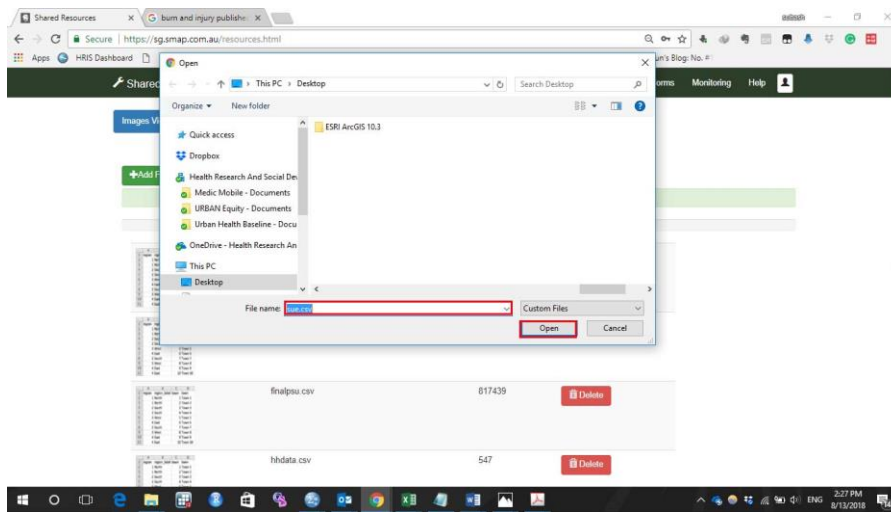
- 4) Click on **“Shared Resources”**.



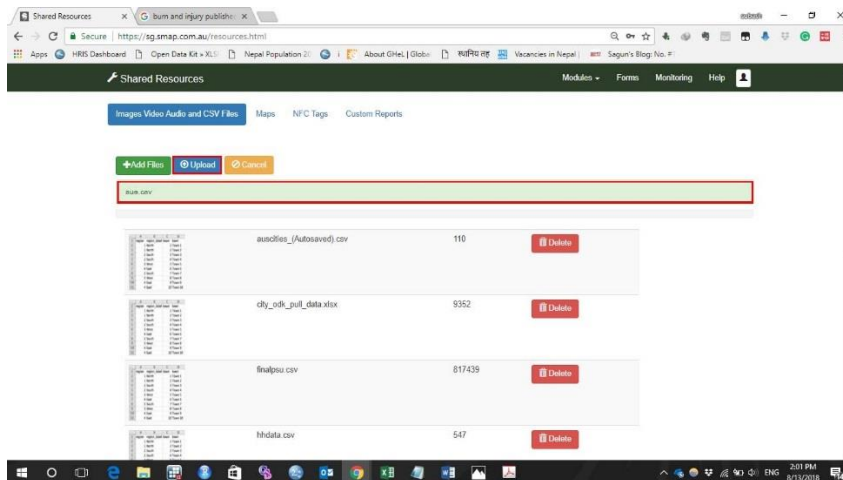
- 5) Then click green color button **“Add files”**.



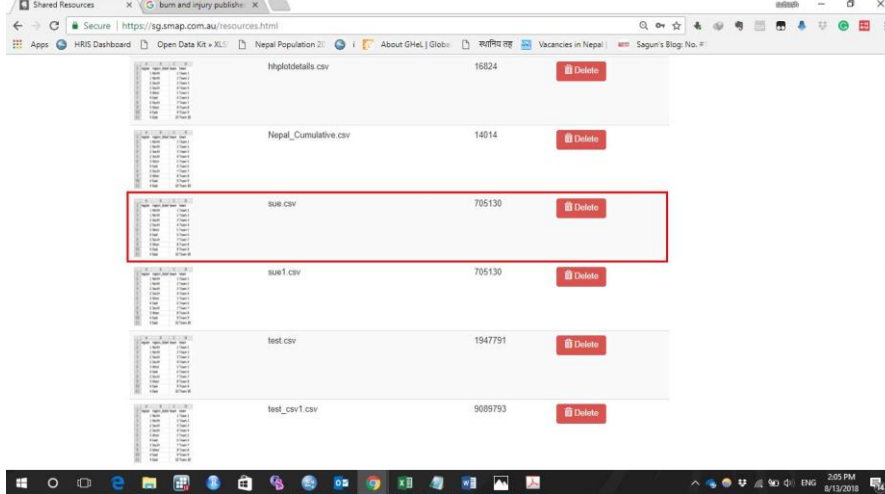
- 6) Folder will appear, then select the **“.csv”** file from the folder.
- 7) After choosing the file, the file name will appear in the **“File name”** as shown in the red box.
- 8) Click **“Open”**.



- 9) When the .csv file will be uploaded, then click **“Upload”**. If .csv file does not have any issues it will upload successfully. The green pop-up notification will appear, showing the success message after the .csv file get upload.

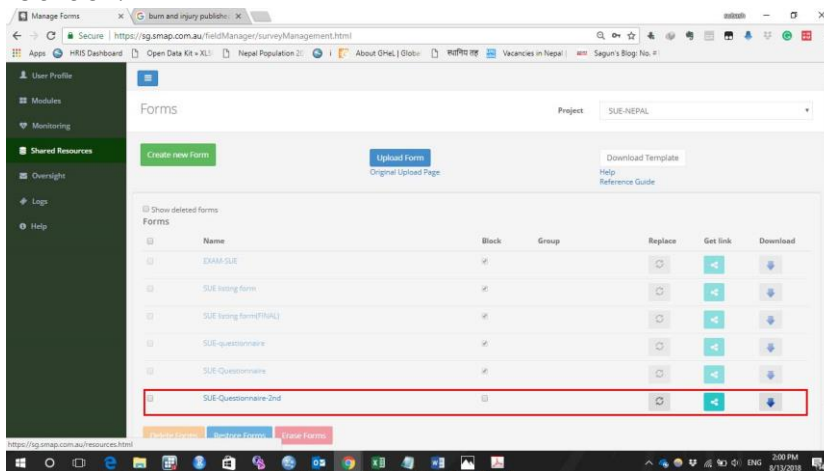


10) The uploaded .csv file will appear as shown in the red box in the photo.



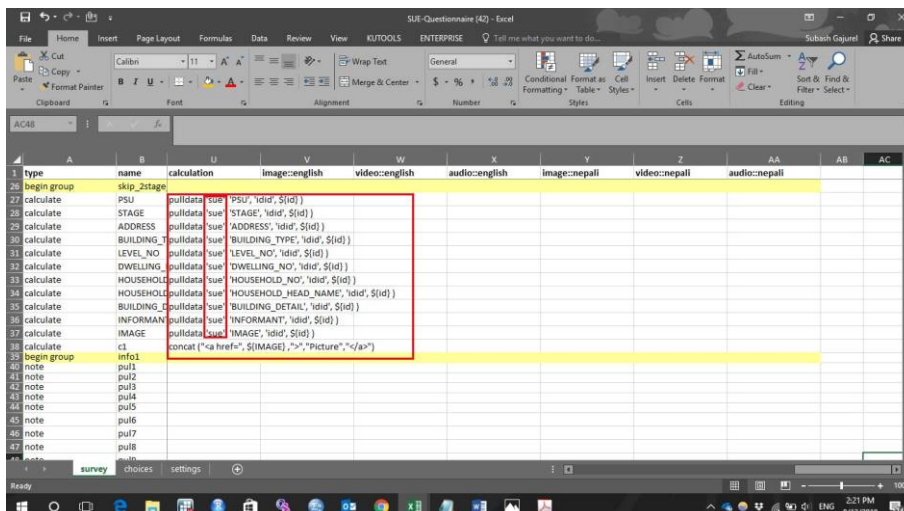
11) Upload the main file. For this you can follow the previous manual file "Manual for uploading forms in sg.smap.com.au".

12) When the file will be uploaded it will appear in the form section as shown in red box.



13) For linking the .csv file with main file, the code should be written as shown in the red box.

14) This code should be placed in the calculation section in main file.



- 15) Red box in row are the variable link with main file to pull the “information”.
- 16) Red box in column are the “ID” through which the information of specific respondent will display.

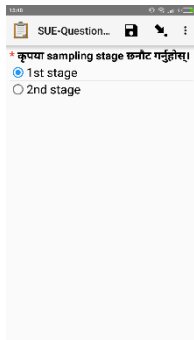
The screenshot shows an Excel spreadsheet with the following data:

Unique-ID	id	Sampled I N.S.N.	SN	prkey	enu	ei(psu)	PSU	STAGE	ADDRESS	BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFORMA	IMAGE	arc	ward
2-1-913	21913	1	1072	5074	913	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Chi	13	
2-2-918	22918	2	1078	5092	918	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Chi	13	
2-3-924	23924	3	1096	5179	924	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Chi	13	
2-4-1173	241173	4	1298	6488	1173	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-5-1175	251175	5	1300	6492	1175	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-6-1178	261178	6	1304	6500	1178	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-7-1187	271187	7	1313	6518	1187	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-8-1197	281197	8	1320	6535	1197	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-9-1253	291253	9	1458	7129	1253	3 Bi	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-10-1269	2101269	10	1468	7155	1269	3 Bi	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-11-1274	2111274	11	1469	7163	1274	3 Bi	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-12-1275	2121275	12	1470	7165	1275	3 Bi	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-13-1277	2131277	13	1472	7189	1277	3 Bi	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-14-1284	2141284	14	1474	7194	1284	3 Bi	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-15-1285	2151285	15	1475	7186	1285	3 Bi	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-16-1290	2161290	16	1478	7194	1290	3 Bi	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-17-1535	2171535	17	1936	9248	1535	3 Bi	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-18-1543	2181543	18	1941	9260	1543	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-19-1546	2191546	19	1944	9264	1546	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
2-20-1556	2201556	20	1952	9285	1556	2 A	2 PSU = 2	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	3	
5/1/2473	513473	1	4340	19813	3473	3 Bi	5 PSU = 5	STAGE = 2	Address = BUILDING LEVEL	NO DWELLING	HOUSEHO	HOUSEHO	BUILDING	INFO = R	"https://s	Bal	5	

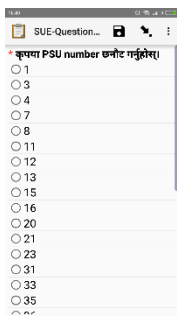
How To: Complete one-stage interview questionnaire

Step 1: Select the stage as one.

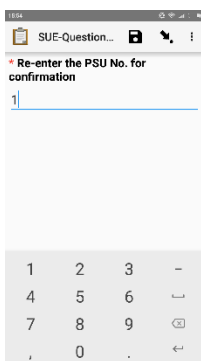
- a. 1st stage (Choose)
- b. 2nd stage



Step 2: List of all PSUs under one-stage will be displayed in list menu. Choose the PSU.



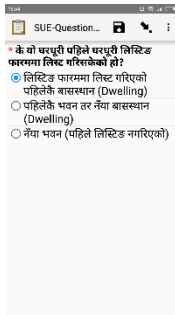
Step 3: Again enter the PSU number. Program will crosscheck step 2 and 3.



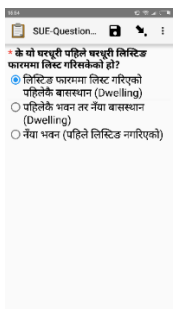
Step 4: If PSU number in step 2 and 3 is same: a question will appear 'Is the dwelling already listed in the dwelling listing form?'

Choose one of the option in the question:

- A) Same as listed in the dwelling listing form, or
- B) Same building but new level/dwelling., or
- C) New building



Step 5: If the dwelling is already listed in the dwelling listing form, click A) Same as listed in the dwelling listing form.



Step 6: After clicking A, Unique dwelling ID will be listed on tablet of the respective PSU. With the help of dwelling list (hard copy) and paper map, choose the correct Unique dwelling ID.

Unique dwelling number is combination of PSU + S.N.+ Building number + Level Number + Dwelling Number.

Ex. 1.20.1565 is unique dwelling ID. S.N. is not more than 30 in each PSU.

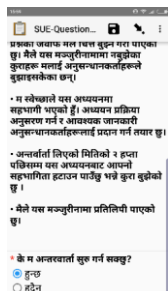
PSU	S.N.	Building #
1	20	1565



Step 7: After choosing the Unique dwelling ID, again enter the same Unique dwelling ID. Address of the dwelling, building detail and picture of the building will be displayed. Please see the information and confirm.



Step 8: Approach to the household and use script and talk with the household member.

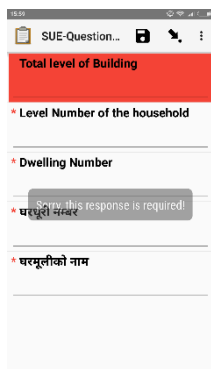


Step 9: Enter level number, dwelling number, household number and name of the household head. In addition, also enter the type of the building. For each household in the dwelling, you should start from the beginning (step 1) and enter the unique dwelling ID.

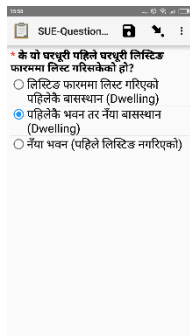


Step 10: After entering all the above-mentioned information, program will automatically check the level number and dwelling number you entered in step 9 with Unique dwelling ID. If all the information are matched, you will directly proceed to the main questionnaire.

If the information are not match, you are asked to correct it. After correction you will proceed to main questionnaire.



Step 11: For new level/dwelling but same building, choose B in step 4.

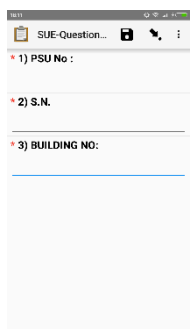


Step 12: For new level/dwelling you need to develop new Unique dwelling ID by yourself. Unique dwelling ID is combination of PSU + S.N.+ Building number + Level Number + Dwelling Number. For new level/dwelling start the S.N from 50 to develop a new Unique dwelling ID. However, make sure that the building number is same.

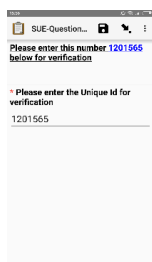
Ex. PSU =32, S.N.=50, building number=15

Unique number = 32.50.15

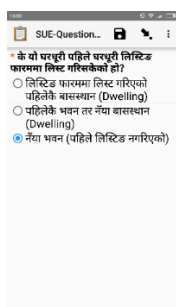
PSU	S.N.	Building #
32	50	15



Step 13: Enter the Unique dwelling ID twice. In addition, also update the Dwelling list (hard copy). Follow steps 8 to 10.



Step 14: For new building, choose C in step 4.



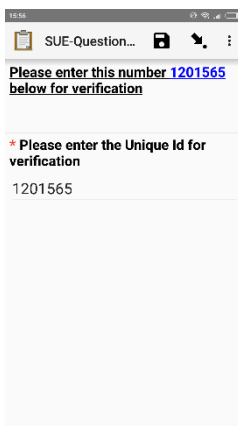
Step 15: For new building, again you need to develop new Unique building ID by yourself. Unique dwelling ID is combination of PSU + S.N.+ Building number + Level Number + Dwelling Number. For new building start S.N from 50 and building number from 1 to develop a new Unique dwelling ID. Here, you are asked to enter interviewer ID too (all the enumerators are assigned with ID). Interviewer ID is entered to avoid duplication.

Ex. PSU =32, S.N.=50, building number=1, Level number=1, Dwelling Number =2 and interviewer ID=12
 Unique number = 32.50.15

PSU	S.N.	Building #
32	50	1



Step 16: Enter the Unique dwelling ID twice. In addition, also update the Dwelling list (hard copy). Follow steps 8 to 10.

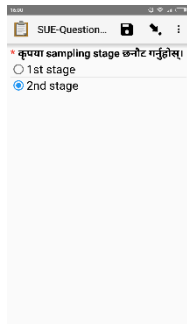


How To: Complete two-stage interview questionnaire

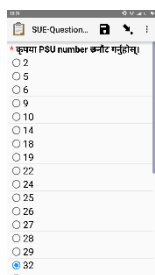
Before the interview, identify the household with help of the household listing in the respective PSU.

Step 1: Select the stage as two.

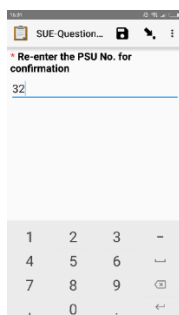
- a. 1st stage
- b. 2nd stage (Choose)



Step 2: List of all PSUs under two-stage will be displayed in list menu. Choose the PSU.



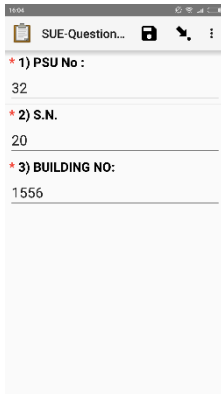
Step 3: Again enter the PSU number. Program will crosscheck step 2 and 3.



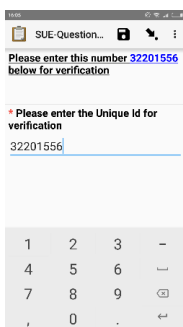
Step 4: After entering the PSU number, all the sampled Unique household ID in the respective PSU will appear in the dropdown menu. Select the Household unique ID where you are going to interview from the drop down menu with the help of Household list (hard copy and paper map).

Eg. 32.20.1556

PSU	S.N.	Building Number
32	20	1556



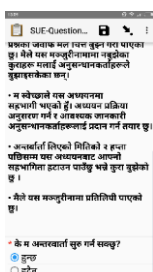
Step 5: After selecting the Unique household ID, enter the Unique household ID again.



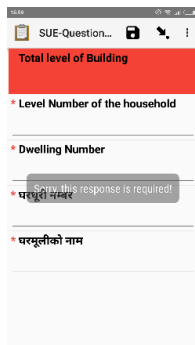
Step 6: If step 3 and 4 are matched; Area name, ward number, level, dwelling, type of building, building detail, building picture and name of the household head will appear in the screen.



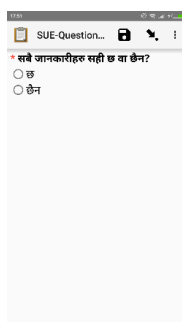
Step 7: Approach to the household and use script and talk with the household member.



Step 8: Please check the information displayed (step 6). A question will appear 'Is this information correct?'



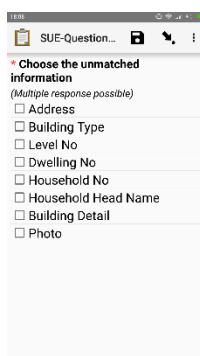
Step 9: If the information is correct, click Yes, else No.



Step 10: After clicking Yes, you will proceed to the main questionnaire.



Step 11: If the information is incorrect, click No. After clicking No, a multiple choice question will appear 'which information is incorrect?' The options will be Area name, Ward number, level, dwelling and type of building. You should choose the option, which is incorrect, and correct it. After correction, you will proceed to main questionnaire.



Appendix D: Exercise Answers

1. Introduction

1. Vertical listing
2. GeoODK
3. c

8. Logistics: fieldwork

1. d
2. c
3. c

2. Definitions

1. b
2. d
3. b
4. b
5. building, level, dwelling,
household

9. Logistics: post-field

1. c
2. a

3. Workflow

1. b
2. d
3. b

4. Survey design and selection

1. b
2. c
3. a

5. Infrastructure and tools

1. b
2. c

6. Logistics: staff

1. a
2. b
3. a

7. Logistics: pre-field

1. a
2. e
3. c