INTERNATIONAL TRAINING ON TOPONYMY

MANUAL GUIDES

19 - 23 JUNE 2023

BALI, INDONESIA

BADAN INFORMASI
GEOSPASIAL

MANUAL GUIDE

Data Collection using KoboToolbox

1.1 Introduction

Form is a fundamental aspect of data collection. It determines the accuracy, quality, and usability of the collected data. In this regard, KoboToolbox may provide valuable assistance in building a well designed form. Kobotoolbox offers a wide range of question types, skip logic, to validation rules and allow users to create forms that are built to user specific data collection needs.

1.2 Requirements

Before getting started with KoboToolbox, it is important to consider the following requirements:

1. Mobile app is only available for Android. There are two choices: ODK Collect and KoboCollect. The choice between the two depends on your personal preference. Both apps provide similar functions for collecting data using forms created in KoboToolbox.

2. The Iphone user is recommended to use BlueStacks to install KoboCollect or may try to use GIC Collect.
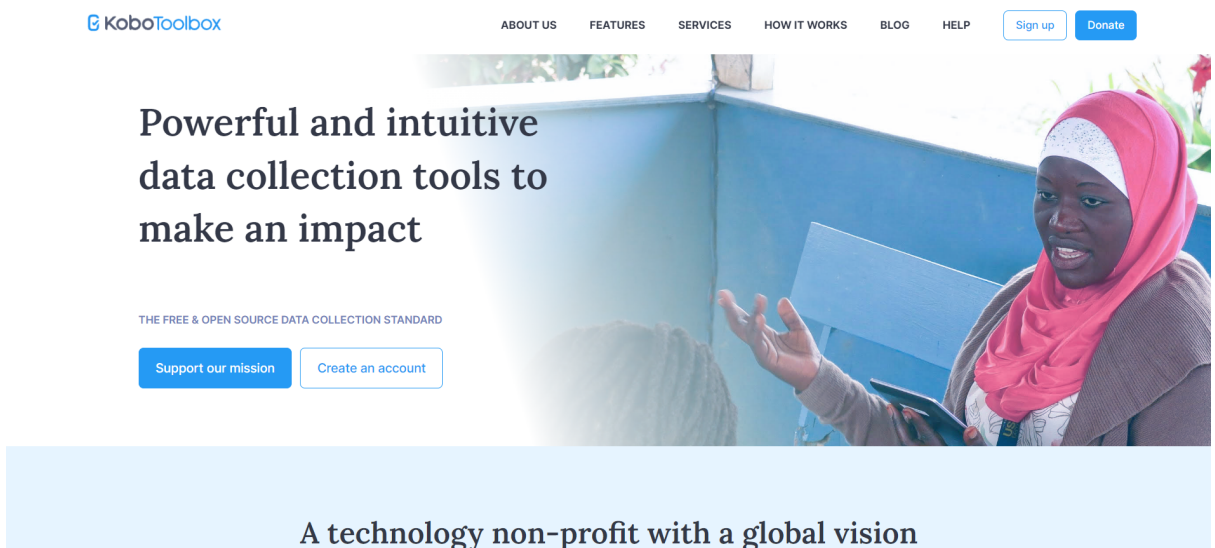
1.3 Installation

The installation of KoboToolbox involves creating an account, installing the necessary mobile app, and setting up a mobile app to connect with the KoboToolbox Account.

1. **Create KoboToolbox Account**
    a. Account Setup
    1) Visit the KoboToolbox website at https://www.kobotoolbox.org/.
    2) Click on the "Sign Up" or "Create an Account".

3) Click on the "Create an Account".

4) Fill in the required information such as your name, organization name, username, email address, password.

5) Complete the registration process by following the instructions provided in the website.
6) Get started to New Project.

2. **Android App Installation - Data Collection App**
   a. On your Android device, open the Google Play Store and search for "ODK Collect" or "KoboCollect" in the search bar.
   b. Install either the ODK Collect or KoboCollect app.
   c. Launch the app once it's installed.

3. **Configure the Android App**
   a. Open the app and go to the "Manually enter project details".
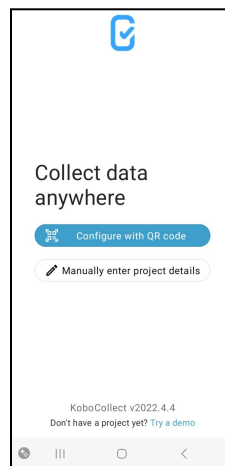
   b. Enter the URL 'https://kc.kobotoolbox.org" and your KoboToolbox account credentials.
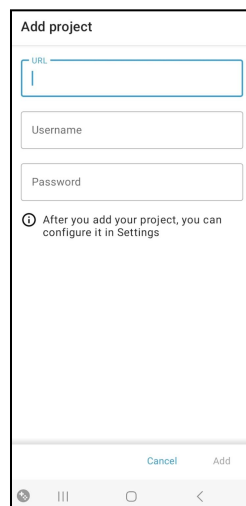
   c. Save the settings and ensure that the app is connected to your KoboToolbox account.
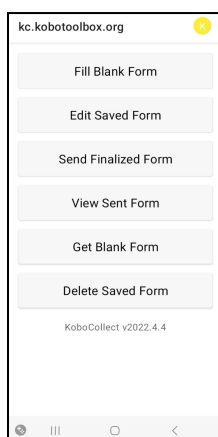
1.4 Create a Survey Form

There are 4 options for creating new forms which are 'Build from scratch', 'Use a template', 'Upload an XLSForm', and 'Import an XLSForm via URL'. You can learn more about creating new forms here: https://support.kobotoolbox.org/new_form.html.

The form we are about to build has 20 questions with various question types. The detailed information of each question can be seen in the table below.

No.	Question	Question Type	Column Name	Mandatory	Domain
1	Type of Geometry	Select One	geom_type	Yes	Point, Line, Area
2	Point	Point	point	Yes	
3	Line	Line	line	Yes	
4	Area	Area	area	Yes	
5	Feature Class	Select One	fclass	Yes	Administrative Boundary, Hydrography, Leisure Area, Populated Place, Railway,Road, Building, Topography, Undersea, Vegetation

UNGEGN | United Nations Group of Experts on Geographical Names

BADAN INFORMASI GEOSPASIAL

|					
6	Feature Type	Select One	ftype	Yes	Administrative Boundary: Administrative Division Hydrography: Bay, Dock(S), Lakes, Ocean, Ponds Leisure Area: Business Center, Military Base, Mining Area, Park, Port Populated Place: Farm Village, Populated Places, Populated Place, Abandoned Populated Place, Religious Populated Place Railway, Road: Oil Pipeline, Road, Railroad, Tunnel, Street. Building: Airport, Cemetery, Castle, Religious Center, Historical Site, Palace, Pagoda, Religious Site, Ruin(S), Shrine, Temple(S), Restaurant, Building(S) Topography: Beach, Cape, Desert, Hill, Mountain Undersea: Bank, Seamount, Basin, Moat, Trench. Vegetation: Bush(Es), Forest(S), Grassland, Tundra, Vineyard. *The choices shown in this question will depend on the response to the Feature Class question. This is called Cascading Select Question
7	Name	Text	name	Yes	
8	Province/State	Select One	prov_state	Yes	Bali
9	Regency	Select One	regency	Yes	Gianyar, Klungkung, Kota Denpasar City, Badung
10	District	Select One	district	Yes	Gianyar: Blahbatuh, Gianyar, Tampaksiring, Ubud Klungkung: Nusa Penida, Banjarangkan, Klungkung, Dawan

|					Denpasar City: Denpasar Selatan, Denpasar Timur, Denpasar Barat, Denpasar Utara
					Badung: Kuta, Mengwi, Abiansemal, Petang, Kuta Selatan, Kuta Utara
11	Province/ State	Text	prov_state	No	
12	Variant Names	Text	var_name	Yes	
13	Previous Name	Text	prev_name	No	
14	Naming History	Text	hist_name	Yes	
15	Language Origin	Text	lang_ori	No	
16	Name Meaning	Text	mean_name	No	
17	Photo 1	Photo	photo_1	Yes	
18	Photo 2	Photo	photo_2	Yes	
19	Voice Recording	Audio	rec	No	
20	Remark	Text	remark	Yes	
21	Survey Date	Date	survey_date	No	

22	Surveyor Name	Text	surveyor	Yes	

In this tutorial, we will utilize the 'Upload an XLSForm' method. We opt for this method due to its extensive flexibility and a wide range of features available within XLSForm. With XLSForm, we can collaborate with other form makers to make the form by making the spreadsheet online to customize various aspects of the form, such as setting up pop up messages when surveyor skip mandatory question, simplifying the process of handling multiple choices questions with numerous options, and implementing cascading select questions.

1. Set up the XLSForm.
    a. Create a workbook. The workbook can be created by using any spreadsheet program.
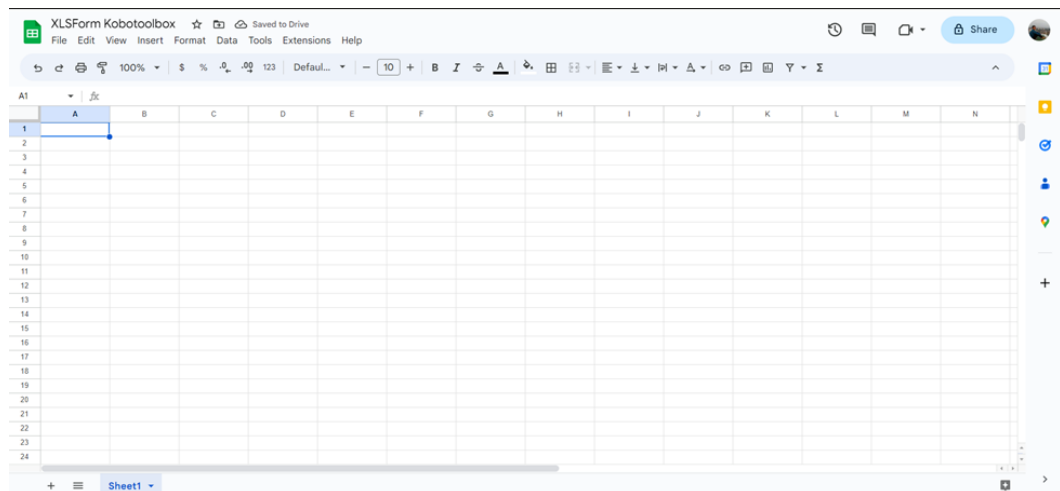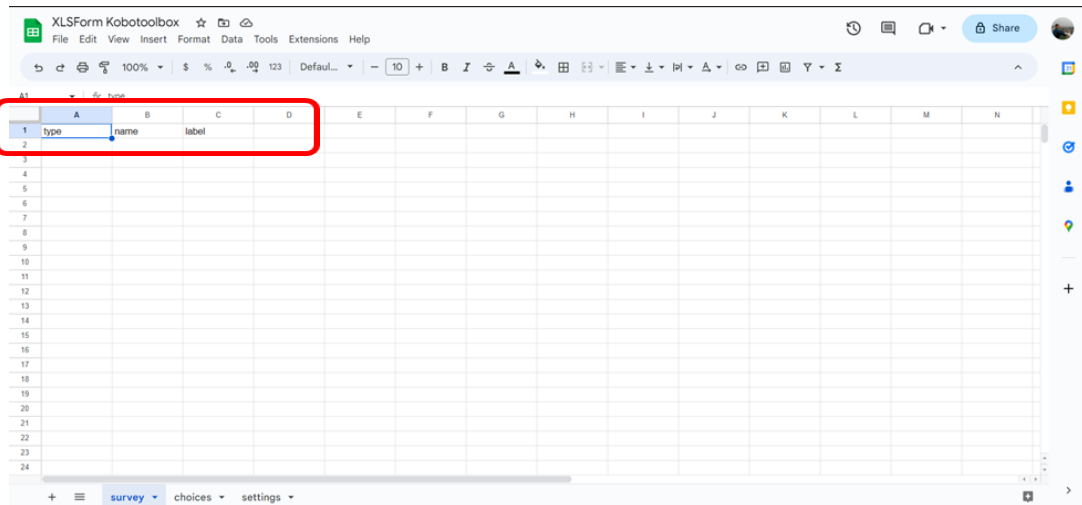

    b. Create three worksheets, namely 'survey', 'choices', and 'settings'. The name of the sheets must be in lowercase.
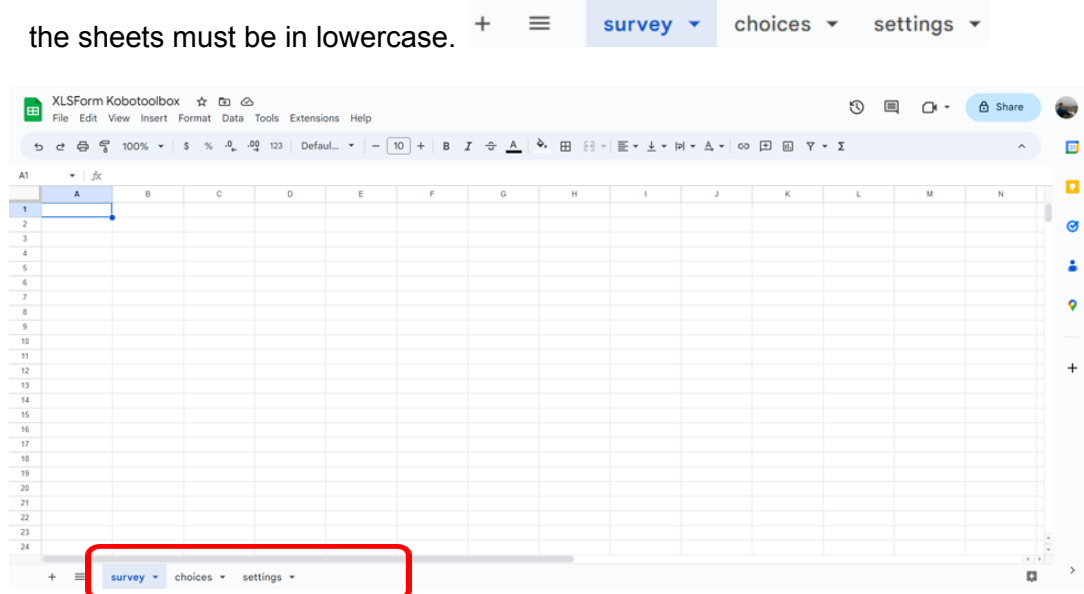
c. In the "survey" worksheet, create three columns with headings 'type', 'name', and 'label'. These three columns are mandatory to build a form and needs. Again, these headers must be in lowercase.

- The 'type' column indicates the question type to each question. There are 26 question types known in XLSForm, but we are not going to use all of them in this form.
- The 'name' column is designated for entering the variable name of each question. Variable names must be unique for each question and cannot symbols (except the underscore). Names are case-sensitive.
- The 'label' column contains the actual text we see for the question in the form.

d. After that, we will add eight more columns: 'hint', 'required', 'required_message', 'relevant', 'choice_filter', 'constraint', 'constraint_message', and 'parameters'. These eight columns are optional for form building. These columns provide features that enhance the clarity and conciseness of the form. However, in this tutorial, we consider them necessary to leverage the full potential of the form.

- The 'hint' column is used to provide additional guidance or instruction related to a question, aiding surveyors with information on how to answer the question.
- The 'required' column enables us to indicate whether a question must be answered, preventing surveyors from skipping it. The value 'yes' is assigned to mark questions as mandatory, while 'no' is used for optional questions. It is mandatory to provide a value in this column for each question.
- The 'required_message' column is used to customize the message displayed to surveyors when they leave a required question unanswered. The
- The 'relevant' is used to apply logical conditions where surveyors can skip questions or display additional questions based on the response from the previous question.
- The 'choice_filter' is used to create a cascading select question. This tool optimizes the efficiency of the form by filtering the choices available in the subsequent questions based on the choices surveyors entered in the

UNGEGN | United Nations Group of Experts on Geographical Names

BADAN INFORMASI GEOSPASIAL

previous question. By employing adaptive choices, it significantly shortens the form length.

- The 'constraint' column allows us to impose limitations on responses. By defining a validation code, answers can be automatically checked whether the response meets the specified criteria we have set up before. Although this feature is commonly applied to numeric questions, in this form, we will utilize constraints for text question type.
- The 'constraint_message' column enables us to leave a message to the surveyor when their answer fails to meet the specified criteria.
- The 'parameters' is used in the image question type to ensure that the submitted image is not too large.

e. The first question in the form will inquire about the type of geometry. In the XLSForm, please write the following entries in their respective cells, precisely as shown in the example

	A	B	C	D	E	F	G	H	I	J	K
1	type	name	label	hint	required	required_message	relevant	choice_filter	constraint	constraint_message	parameters
2	select_one geomtype	geom_type	Type of Geometry	Select the appropriate geometry type that can represent the selected geographical features	TRUE	Oops! This question is mandatory. Please provide a response before moving to the next question.					

- In the 'type' column, please write 'select_one geomtype'. We use 'select_one' to indicate that the question type is a multiple-choice question where only one answer can be selected. Additionally, 'geomtype' will serve as the ID that connects this question to its corresponding choices that will be provided in the 'choices' sheet.
- In the 'name' column, please write 'geom_type'.
- In the 'label' column, please write 'Type of Geometry'.
- In the 'hint' column, please write 'Select the appropriate geometry type that can represent the selected geographical features.
- In the 'required' column, please write 'true'.
- In the 'required_message' column, please write 'Oops! This question is mandatory. Please provide a response before moving to the next question.'
- In the 'relevant', 'choice_filter', 'constraint', 'constraint_message', and 'parameters' columns, please leave it blank.

f. Continue filling the rows following the provided example provided at Example of XLSForm - The International Training on Toponymy.xlsx.

- In the 'Point', 'Line' and 'Area' questions, the 'relevant' column is filled with "${geom_type} = 'point'", "${geom_type} = 'line''', and "${geom_type} = 'area'" respectively. This condition is set so that the respective question will only be displayed when the surveyor selects the corresponding answer in the "Type of Geometry" question. The lowercase format of 'point', 'line' and 'area' in the "${geom_type} = 'point'", "${geom_type} = 'line''', and "${geom_type} = 'area'" is used because the choices defined in the 'name' column of the 'choices' sheet for 'Type of Geometry' are all in lowercase.
- In the 'Feature Type' question, the 'choice_filter' column is filled with "fclasses=${fclass}". This indicates that the available choices of Feature Type question will depend on the selected choice in the 'Feature Class' question. In the expression "fclasses=${fclass}", "fclasses" refers to the header in the

'choices' sheet, and 'fclass' refers to the 'name' in the 'survey' sheet where the 'Feature Class' question is defined.

- The same goes for the 'district' question. In the 'district' question, the 'choice_filter' column is also filled with "regen=${regency}". This indicates that the available choices in the 'district' question will depend on the selected choice in the 'regency' question.

- The expression "regex(., '^[\S\s]{1,254}$')" is used as a constraint in multiple questions within the form. It is to ensure that the response in a text response to a maximum of 254 characters long while allowing any type of character to be entered. We use it in 'Name', 'Variant Names', 'Previous Names', 'Naming History', 'Language Origin', 'Name Meaning', 'Remark', and 'Surveyor Name' questions.

- The 'constraint_message' is utilized only for questions that have a constraint applied. The message is 'We're sorry, but the input you provided has exceeded the maximum character limit of 254. Please enter a value with a maximum of 254 characters.' This message serves as a reminder to surveyors when their input surpasses the specified character limit. By displaying this message, we prompt surveyors to revise their response and ensure it falls within the allowed length.

g. In the 'choices' worksheet, create three columns with headings list_name, name, and label.
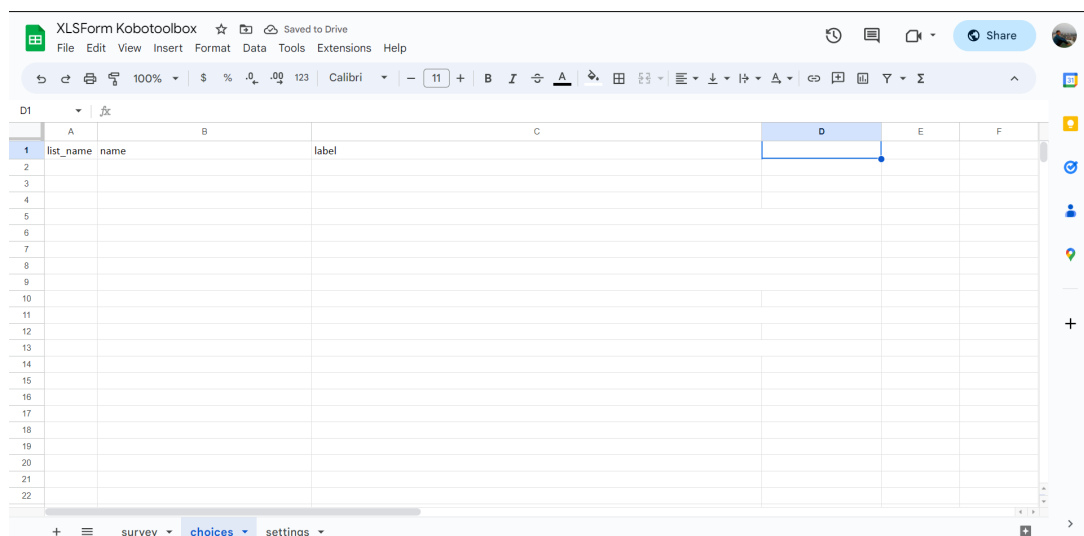
- Create three columns with headings 'list_name', 'name', 'label'. These three columns are mandatory to build a form with multiple question choices. These headers must be in lowercase.

- The 'list_name' column enables us to group a set of related choices together. This column will be filled with the code of a question from the 'survey' sheet, which represents the question to which the choices belong. By using this column, all the choices associated with that question will appear together under that specific question in the form.

- The code of a multiple choice question refers to the 'type' column of the 'survey' sheet. For example, in the 'survey' sheet we have written, one of the questions with multiple choices is 'prov_state'. Its 'type' column is 'select_one

prov'. 'select_one' the question type, while 'prov' is the code of the question that we can refer to in the 'list_name' of 'choices' sheet.
- In the form we have several multiple choice questions which are 'geomtype', 'fclass', 'ftype', and 'prov', 'regency', and 'district'.
- The 'name' column specifies the unique variable name for that particular answer choice.
- The 'label' column is used to specify how the answer choices should be displayed on the form. It represents the text that will be visible to the surveyor.

h. 'ftype' question is a cascading select question. To build this type of question, we need to add an additional column in this sheet. The header of this column refers to the 'choice'_filter of the 'survey' sheet, specifically from the 'ftype' question.

| 7 | select_one ftype | ftype | Feature Type | Please select the feature type that best describes the location you are working with based on the Feature Class you have chosen | TRUE | Oops! This question is mandatory. Please provide a response before moving to the next question. | | fclasses=${fclass} |

In the 'choice_filter' of the 'survey' sheet, it is written 'fclasses=${fclass}'. While 'fclass' refers to the variable name of 'feature class' question, 'fclasses' will be used for the header of the additional column in the 'choices' sheet. This will set up the cascading relationship between 'ftype' question and 'fclass' question.
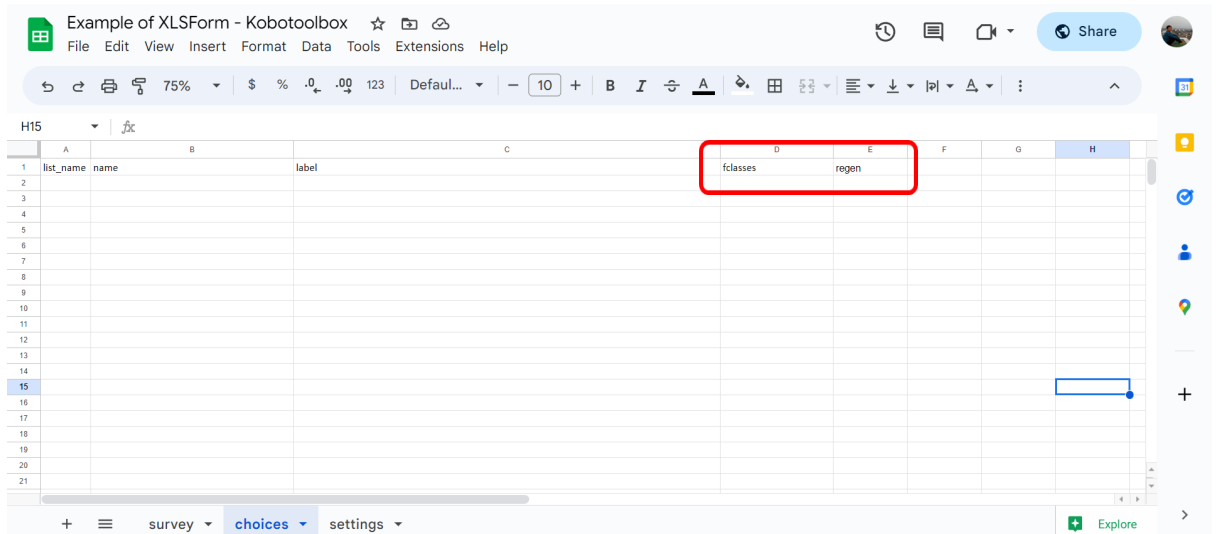

i. Fill in the rows in 'choices' sheet following the provided example provided at Example of XLSForm - The International Training on Toponymy.xlsx.
j. The 'settings' sheet is optional. There are several headers available which are 'form_title', 'form_id', 'version', 'instance_name', 'default_language', 'public_key', 'submission_url', 'style', and 'name'.
   In this form, we use 'form_title', 'form_id', and 'default_language' headers. You can follow the provided example provided at Example of XLSForm - The International Training on Toponymy.xlsx.'
k. You can learn more about creating new forms here: https://xlsform.org/en/.
2. After finishing building the form, download the form in Microsoft Excel format.
3. Access your KoboToolbox account, click 'New' and choose 'Upload an XLSForm'.

UNGEGN | United Nations Group of Experts on Geographical Names

BADAN INFORMASI GEOSPASIAL

4. After you successfully upload the XLSForm in KoboToolbox, you can perform a form check to ensure that it has been created correctly according to your need.

1.5 Survey Procedure

Before collecting data, we need to set up sharing permission in KoboToolbox to allow other users to access and collect data using our form.

1. **Set up sharing permission in KoboToolbox website**
   a. Access your KoboToolbox account through the website and navigate to the project.
   b. Go to the "Summary" section and click "Share Project".
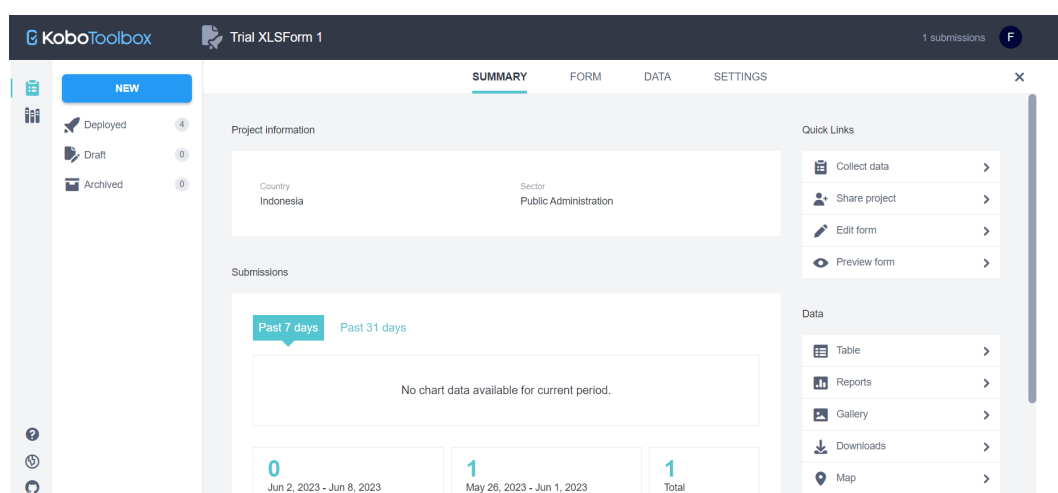
   c. Click the "Add User" and write the username of the targetted users and set up their authorization level accordingly.
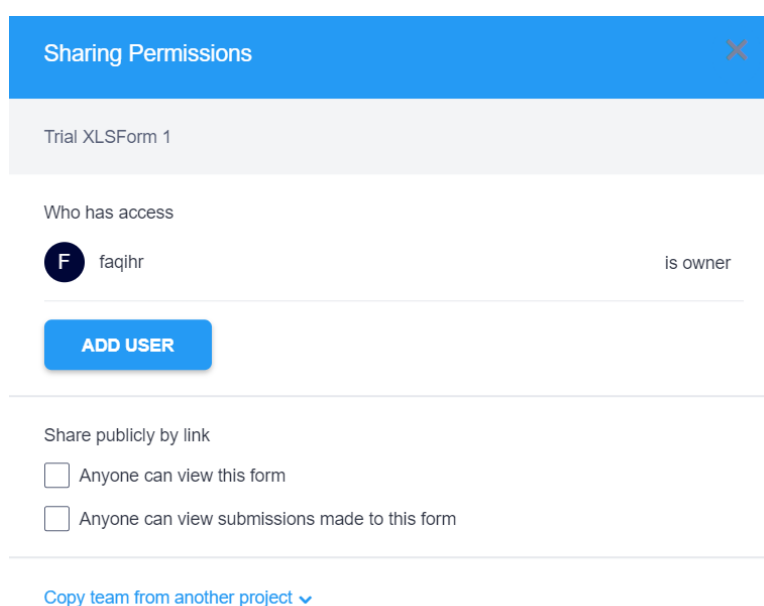
   d. Go to the "Form", and choose the appropriate data collection method for our form. KoboToolbox offers several options such as online or offline with multiple

submission to online only but one submission per respondent and use Android Application.
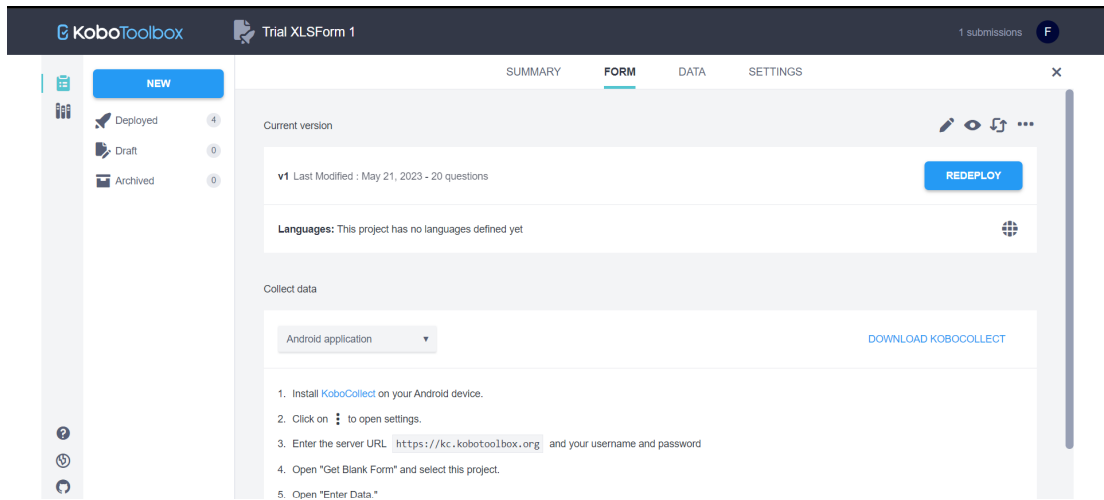
By selecting the appropriate data collection method, we can define how users will access and submit data using our form. In this module, we will focus on Data Collection Using Android Application.
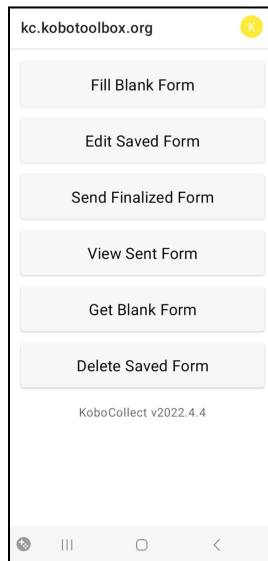
e. Clicking on the "Deploy" button in the "Form" section will update the form and make it available for users who want to collect data. It is important to remember that you should click "Deploy" every time you make changes to the form in order for the updates to take effect.

UNGEGN | United Nations Group of Experts on Geographical Names

BADAN INFORMASI GEOSPASIAL
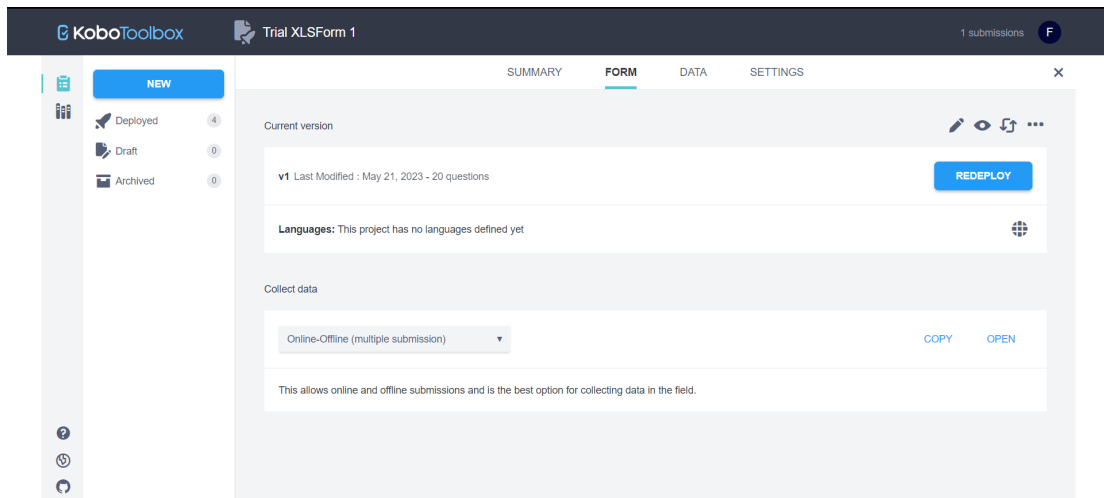
2. Data Collection in Android Apps
   a. Download Forms
      1) Tap on "Get Blank Form".

2) KoboCollect will sync with your KoboToolbox account and download the available forms associated with your account. You can see several information such as the version and the ID of the Form.
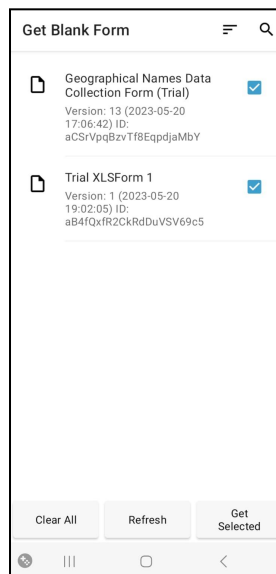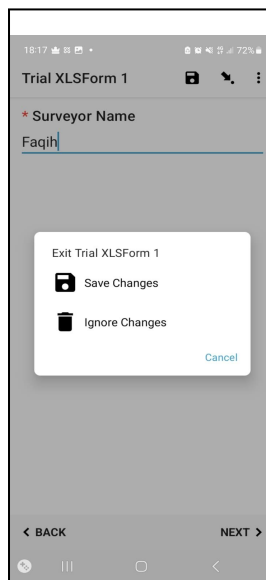

3) If your form you need does not appear or not the newest version, tap on "Refresh" button to update the list of available forms.
4) From the list, select the form you want to use for data collection and tap "Get Selected".
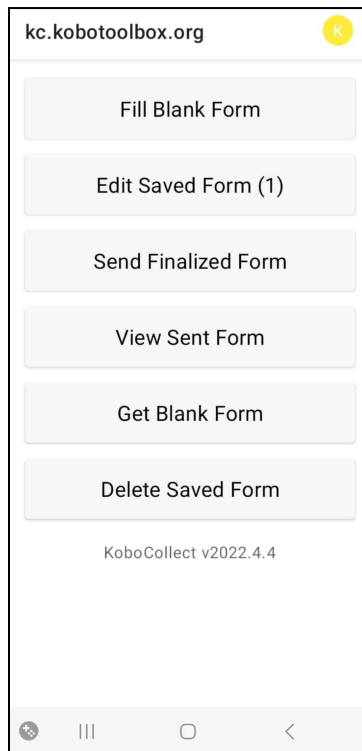
b. Fill out the Form
   1) Tap on "Fill Blank Form" and choose the form you want to fill.
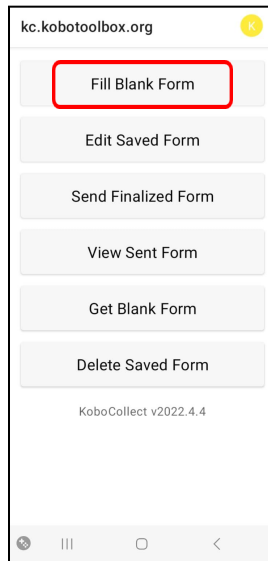
   2) Start filling out the form by entering the required information.
   3) Depending on the form design, you may encounter skip, validation checks and other features implemented in the form.
   4) You can pause the process of filling out the form or discard the form by tap go back on your phone. If you want pause the process, you can also tap directly on 💾 .
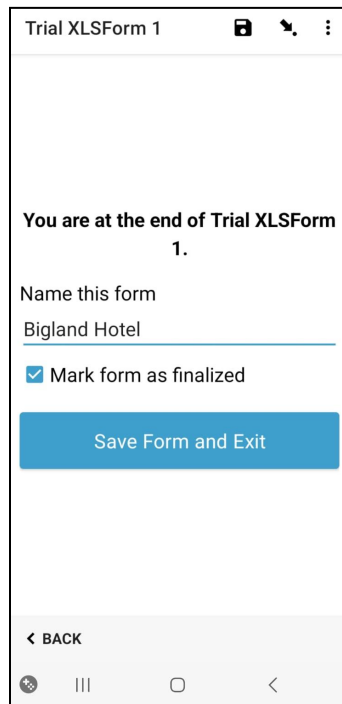

   5) If you choose to save changes, you can get back to the form by clicking the "Edit Saved Form (1)".

UNGEGN | United Nations Group of Experts on Geographical Names

BADAN INFORMASI GEOSPASIAL

6) You can click on "Go To Prompt" ⬩ in the Android app to navigate directly to a specific question within the form. This feature allows you to quickly jump to a particular question during data collection.
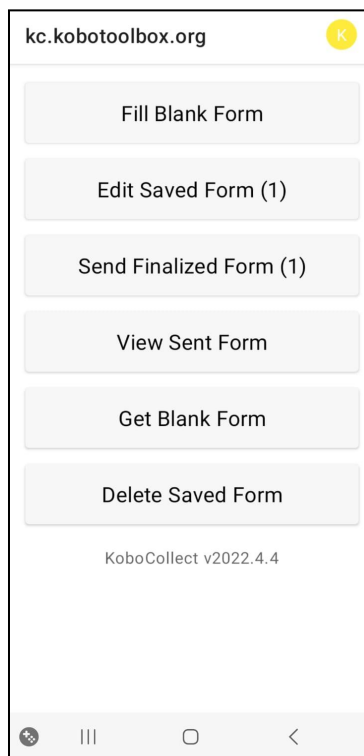
UNGEGN | United Nations Group of Experts on Geographical Names

BADAN INFORMASI GEOSPASIAL
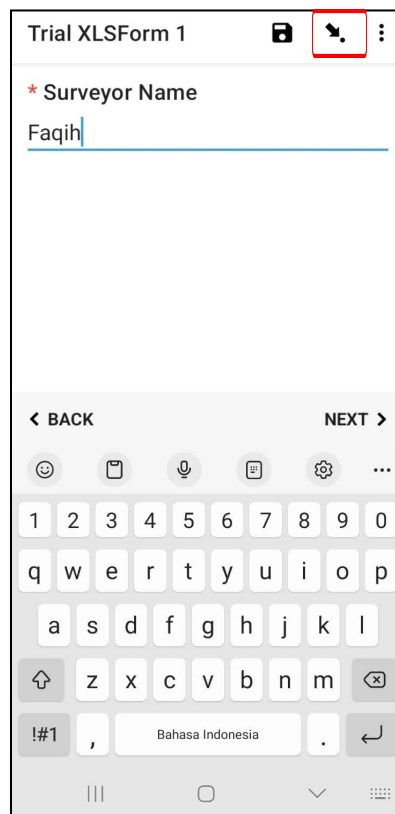
7) Once you have completed filling out the form, tap on the "Save Form and Exit".
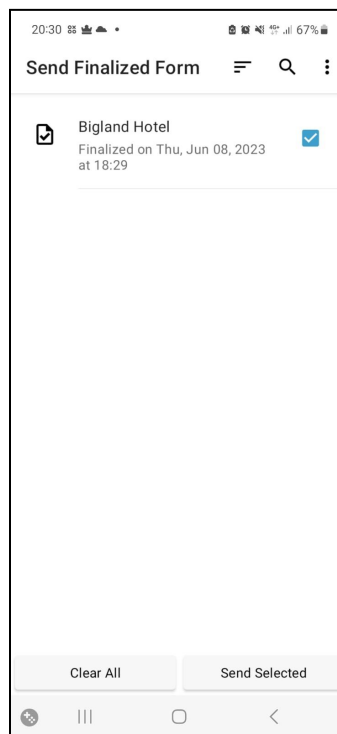

8) You can review and edit your answer on the form by tapping "Edit Saved Form" in the main menu.

c. Submit Data
   1) Tap on "Send Finalized Form".

2) Select the form you want to submit and tap "Send Selected".

1.6 In-app Data Processing and Data Extraction

Data processing and data extraction can be accessed through the "Data" section of the respective form. Here, you can perform various actions on the collected data, such as viewing, validating the data, editing, analyzing, and exporting data. There are "Table", "Reports", "Gallery", "Downloads", and "Map" features.

1. "Table" feature displays collected data in a tabular format. This allows us to view, filter, sort, edit, delete, or validate individual submission. There are three validation status we can choose, "Not Approved", "Approved", and "On Hold". This feature helps in distinguishing between reviewed and pending submissions, ensuring data quality and accuracy.

2. "Reports" feature allows us to generate customized reports based on collected data. We can create visualizations, charts, and graphs to gain insight and present our findings effectively. The report can be customized according to our need from applying filters, grouping, choosing visualization color to exporting and sharing reports.

3. "Gallery" feature shows the media files collected during data collection. It allows us to view and manage the media files. This feature can also sort the media files by specific questions or date and time of submission.

4. "Downloads" feature enables us to export data in various file formats, such as XLS, CSV, GeoJSON.

- "Select Export Type" enables us to choose the desired file format for exporting your data. The options include XLS (Excel), CSV (Comma-Separated Values), GeoJSON, Media Attachments (ZIP), GPS Coordinates (KML).
- "Value and Header Format" allows us to choose between "Labels" and "XML values and headers". "Labels" will make exported files use the question label as column headers and value labels for the responses. Meanwhile, "XML values and headers" will set column headers and responses using variable names and XML values. By choosing "XML values and headers", it ensures that the exported data maintains the original structure of the form.
- "Export Select Many questions as…." allows us to choose how to export the responses from questions where multiple answers can be selected. "Single and separate columns" is chosen because this is the default and there are no "Select Many" questions in the form.
- "Include data from all version" feature allows us to download the data from all of our form versions. When it's unchecked, the downloaded data will only be from the latest deployed form version.
- "Include media URLs" option allows us to include the URLs of media files in the exported data. By Including media URLs, we can easily retrieve the media files associated with each submission and view or use them as needed.