



FEED THE FUTURE

The U.S. Government's Global Hunger & Food Security Initiative



Capacity Building Training and Dissemination on Seed Demand Data Management for Staff at different Levels in Arsi Zone *Final Report*



USAID
FROM THE AMERICAN PEOPLE



Activity Title: Feed the Future Global Supporting Seed Systems for Development activity.

Activity start date and end date: Aug 24, 2018 – Dec 22, 2023

Cooperative agreement number: 7200AA18LE00004

Document title: Capacity Building Training and Dissemination on Seed Demand Data Management for Staff at different Levels in Arsi Zone
Final Report

Publication date : November 2023

Authors: HEDBEZ Business and Consultancy, PLC.

Sponsoring USAID office: LOC Unit, Federal Center Plaza (SA-44)/M/CFO/CMP

Technical office: USAID/RFS/CA

AOR name: Daniel Thomson

Activity Goal: Improved functioning of the high-impact integrated seed systems

Language of document: English

Submitted on behalf of: Catholic Relief Services

Submitted by: Jason Sullivan, Chief of Party S34D activity
Catholic Relief Services
228 West Lexington Street, Baltimore, MD 21201
Jason.Sullivan@crs.org

Cover page photo credit: Bezabih Emana

DISCLAIMER

This report was made possible by the generous support from the American people through the U.S. Government's Feed the Future initiative and the United States Agency for International Development through Cooperative Agreement 7200AA18LE00004. The contents are the responsibility of Catholic Relief Services and do not necessarily reflect the views of USAID or the United States Government.

Feed the Future Consortium Partners in the Feed the Future Global Supporting Seed Systems for Development activity:



Contents

| | |
|--|----|
| 1. Introduction..... | 5 |
| 2. Objectives of the Training Workshop..... | 5 |
| 3. Approach | 5 |
| 4. The Training Workshop Program..... | 6 |
| 5. Training Workshop Participants | 6 |
| 6. Content of the Training Materials/Workshop Documents..... | 7 |
| 7. Evaluation of the Training Workshop and Outputs..... | 8 |
| 7.1 Evaluation of the training | 8 |
| 7.2 Evaluation of the output of the capacity building support..... | 8 |
| Annexes | 10 |
| Annex 1: List of Training Workshop Participants..... | 10 |
| Annex 2: Training/Workshop Manual..... | 11 |
| Introduction..... | 12 |
| Targets of the Manual | 12 |
| Training Workshop Program | 12 |
| Objectives of the Training Workshop..... | 13 |
| Approach..... | 13 |
| Overview of the Seed Demand Data Collection Instruments..... | 13 |
| Data Cleaning | 14 |
| Common Errors in Data..... | 14 |
| Identifying Errors | 15 |
| Measures to Clean Data | 18 |
| Data Quality Assurance | 20 |
| Data Aggregation/Summary | 20 |
| Augmented data: | 22 |
| Reporting..... | 24 |
| Annex 3: Training Evaluation Form..... | 28 |

1. INTRODUCTION

S34D conducted a study on seed demand forecasting in Ethiopia in 2020 and provided recommendations to improve seed demand data collection and reporting. Hence, piloting the recommendations on seed demand forecasting is implemented in a sample of 10 kebeles in 4 woredas of Arsi zone in 2023. For this purpose, detailed survey questionnaire was prepared, 10 Das (1 from each kebele) were trained on the content of the questionnaire and technique of using Tablet for data collection. The training on seed demand data collection was conducted during May 18 to 23, 2023 at Assela, in the Arsi capital zone. Following the training, seed demand data were collected by the trained DAs using the Tablet provided by CRS, as part of kebele level capacity building.

The overview of the questionnaire used for data collection are refreshed and key variables need for analysis are covered in this training. This training aims to build the capacity of the staff of the zone, woreda and kebele as well as Hetosa Union to be able to electronically collect data, clean, analyze and report the data to their respective higher bodies.

2. OBJECTIVES OF THE TRAINING WORKSHOP

The aim of the workshop is creating awareness on the application of the seed demand data collection, data management and reporting and discussing on the ways forward.

This capacity building training enables the trainees to detect errors or inconsistencies in the dataset, clean the data and ensure data quality, analyze the data and report. The specific objectives are to enable the trainees to:

1. check consistencies and inconsistencies in the data
2. conduct data cleaning
3. ensure data quality
4. summarize data and
5. report the data

3. APPROACH

The dissemination activities covered presentation of the approach used to implement the seed demand data collection, the capacity building made and future application of the approach. This was done prior to the training on data management. The participants of the dissemination workshop were also trained on data management and evaluated the training as well as the outputs of the capacity building activities.

The training workshop was conducted in a participatory method:

- Power point presentation of the training manual was provided by the consultants.
- Discussions on issues that need clarification were made.

- Group exercises were done using their own data. Under the supervision and follow up by the consultant, the group members exercised data management activities at different levels:
 - groups for kebele level data management
 - groups for woreda level data management
 - group for zone level data management
- Materials: Flip chart, LCD, note pads and pens were provided
- The outcomes of the group works were discussed in a plenary so that different groups learn from each other.

4. THE TRAINING WORKSHOP PROGRAM

The program started with registration of participants and introduction of each participant: name, where they came from and their responsibility. The training workshop program is given in Table 1.

Table 1: The training workshop program

| Date | Time | Activity | Responsible person |
|-----------------------------|-------------|--|---------------------------|
| Friday Oct. 27, 2023 | Morning | Desktops transported to Assella | CRS Ethiopia |
| | 13:30-14:30 | Orientation on Desktops and program application | CRS Ethiopia |
| | 15:00-17:30 | Exercising with Excel Application | Dr. Bezabih Emana |
| Saturday Oct.28, 2023 | 08:30-9:00 | Introduction and Objectives of the Training | Dr. Bezabih Emana |
| | 09:00-10:00 | Overview of the seed data collection instruments and key variables for analysis (Usual and Augmenting) | Dr. Bezabih+ Girma Fikadu |
| | 10:00-10:30 | Discussion | Participants |
| | 10:30-11:00 | Refreshment break | HEDBEZ |
| | 11:00-12:30 | Data cleaning | Dr. Bezabih+ Girma Fikadu |
| | 12:30-13:00 | Data consistency | Dr. Bezabih+ Girma Fikadu |
| | 13:00-14:00 | Lunch break | HEDBEZ |
| | 14:00-15:30 | Group exercise: data quality checking and detecting errors | Participants |
| | 15:30-16:00 | Refreshment break | HEDBEZ |
| | 16:00-17:00 | Data cleaning: Fixing errors | Dr. Bezabih+ Girma |
| Sunday Oct. 29, 2023 | 08:30-9:30 | Data aggregation | Dr. Bezabih+ Girma |
| | 09:30-10:30 | Group exercise: data aggregation at different levels | Participants |
| | 10:30-11:00 | Refreshment break | HEDBEZ |
| | 11:00-12:30 | Discussion on aggregation/summary making | Participants/facilitated |
| | 12:30-13:00 | Adjustment of the demand data | Dr. Bezabih+ Girma |
| | 13:00-14:00 | Lunch break | HEDBEZ |
| | 14:00-15:30 | Reporting data | Dr. Bezabih+ Girma |
| | 15:30-16:00 | Refreshment break | HEDBEZ |
| | 16:00-17:00 | Group exercise on data reporting at different levels | Participants |
| | 17:00-17:30 | Ways forward and concluding remark | Ebrahim + Bezabih |

5. TRAINING WORKSHOP PARTICIPANTS

The training workshop participants were development agents in the 10 sample kebeles, who were trained on seed demand data collection and collected the data during May 18 to 23, 2023. In addition, seed experts in the

4 woredas and experts responsible for seed data management at Arsi zone Agriculture Office and Hetosa Farmers' Multipurpose Cooperatives participated in the training.

A total of 17 persons participated in the training conducted at Asela town, the capital of Arsi zone. The training was conducted at the premises of Arsi zone Agriculture in order to use the newly delivered desktops to the sample kebeles, woredas, the zone and the union by S34D as part of the capacity building. Since CRS Ethiopia delivered the desktops prior to the training, the trainees practiced the unpacking, connecting parts of the desktops and using them for the training. At the end, they also practiced the repacking. This would not have been possible if the training was organized outside the zone office premises. The list of the training participants is given in **Annex 1**.



6. CONTENT OF THE TRAINING MATERIALS/WORKSHOP DOCUMENTS

In order to disseminate the activities accomplished by S34D, a PowerPoint presentation of the training manual/workshop document given in **Annex 2** was made by the consultant. The participants were invited to ask questions and to participate in the discussion. The manual was also distributed to all participants. The exercises were supervised.

7. EVALUATION OF THE TRAINING WORKSHOP AND OUTPUTS

7.1 Evaluation of the training

At the end of the training workshop, each of the training participants was given an evaluation form (Annex 3) to evaluate the training. Figure 1 shows that all of the participants agree or strongly agree to the overall training process or content. For some 6% of them, the training content was not easy, and the time given for the training was short.

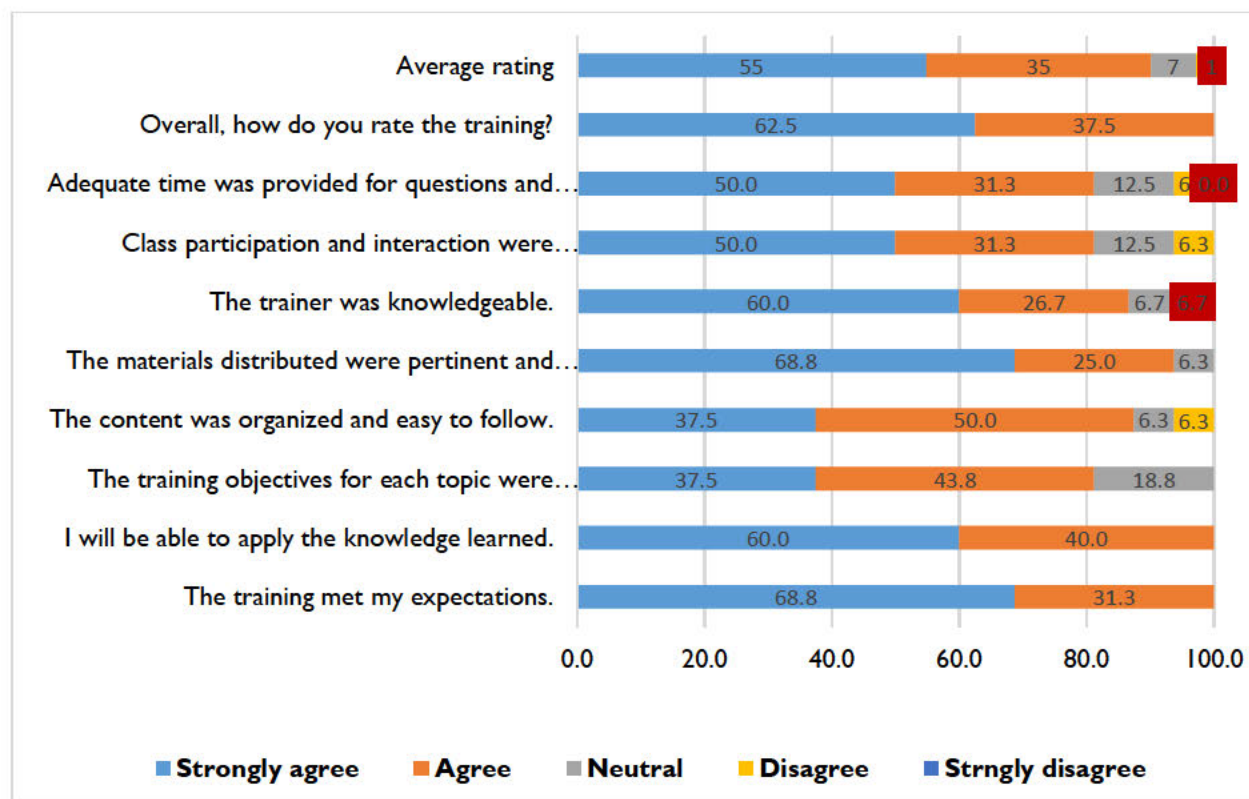


Figure 1: Proportion of training participants rating the training (%)

7.2 Evaluation of the output of the capacity building support

Discussions were held with the training workshop participants coming from the 10 kebeles (DAs), 4 woredas (woreda input experts), representative of Hetosa Cooperative Union and the zonal seed experts of Arsi zone. All of them have high appreciation of the capacity building interventions and considered it as unique compared to what they have experienced over years. The summary of their evaluations include:

- Training on seed data collection using electronically is efficient timewise, so life data can be collected.
- “Provision of Tablets of high quality is a big asset. Compared to the kebeles nearby the sample kebeles selected for this piloting, we have good knowledge of data collection and data management” (DAs’ remark)

- Training on data management is “eye opening”. “We collect data, compile and report to higher offices, (had not) thought about data cleaning” (reported by kebele, woreda and zonal agriculture participants) and data aggregation prior to the training.
- Provision of desktops will surely help us for data management, data and report storage and producing reports.
- We thank and appreciate CRS for supporting us (all say this).

There were also some recommendations given by the participants:

1. The time between data collection and capacity building through Desktop provision and data management training was long. In the future, data management training should be done immediately after the data collection is completed.
2. The scope of the piloting is small: 10 kebeles in 4 woredas. There are several kebeles, where expansion to the kebeles not covered and new woredas in the zone is needed.
3. There is a need for follow up or monitoring of the application of the knowledge gained. This also includes the license renewal of the software of the Desktops.
4. There should also be a printer supplied not just a desktop.

ANNEXES

Annex 1: List of Training Workshop Participants

| No. | Name | Gender | Organization | Position | Email | Telephone |
|-----|------|--------|--|-------------------------------|-------|-----------|
| 1 | | Male | Lode Hetosa woreda, Adda Maree kebele | Development Agent (DA) | | |
| 2 | | Male | Lode Hetosa woreda, Madda Bishaanii kebele | DA | | |
| 3 | | Female | Lode Hetosa woreda, Fursaa kebele | DA | | |
| 4 | | Female | Tiyo woreda, Aboosara Alkoo kebele | DA | | |
| 5 | | Female | Tiyo woreda, Odaa Dhawataa kebele | DA | | |
| 6 | | Male | Hetosa woreda, Seeroo Alkasaa kebele | DA | | |
| 7 | | Male | Hetosa woreda, Gondee Fincamaa kebele | DA | | |
| 8 | | Male | Hetosa woreda, Shaaqii Shararaa kebele | DA | | |
| 9 | | Female | Munessa woerda, Dobaa Ashee kebele | DA | | |
| 10 | | Male | Munessa woerda, Garranboota Lolee kebele | DA | | |
| 11 | | Male | Munessa woreda Agri. Office | Seed expert | | |
| 12 | | Male | Lode Hetosa Agr. Office | Input expert | | |
| 13 | | Male | Hetosa Agr. Office | Input expert | | |
| 14 | | Female | Hetosa Cooperative Union | IT and data management expert | | |
| 15 | | Male | Tiyo Agr. Office | Extension head | | |
| 16 | | Male | Arsi Zone Agri Office | Input expert | | |
| 17 | | Male | Arsi Zone Agri Office | Property expert | | |

Annex 2: Training/Workshop Manual

TABLE OF CONTENTS

| | |
|--|----|
| Introduction..... | 12 |
| Targets of the Manual | 12 |
| Training Workshop Program..... | 12 |
| Objectives of the Training Workshop..... | 13 |
| Approach..... | 13 |
| Overview of the Seed Demand Data Collection Instruments..... | 13 |
| Data Cleaning | 14 |
| Common Errors in Data..... | 14 |
| Identifying Errors | 15 |
| Measures to Clean Data | 18 |
| Data Quality Assurance | 20 |
| Data Aggregation/Summary | 20 |
| Augmented data: | 22 |
| Reporting..... | 24 |

INTRODUCTION

Support piloting seed demand forecasting S34D recommendations in select woredas and kebeles in Arsi zone started with sampling of 10 kebeles in 4 woredas of Arsi zone for collection of seed demand data. For this purpose, detailed survey questionnaire was prepared, one DA from each kebele was trained on the content of the questionnaire and how to use Tablet for data collection. The training was conducted during May 18 to 23, 2023 at Assela. Following the training, seed demand data were collected by the trained DAs using the Tablet provided by CRS.

The overview of the questionnaire will be refreshed, and key variables need for analysis will be covered in this training. The training aims to build the capacity of the staff of the zone, woreda and kebele to be able to electronically collect data, clean, analyze and report the data to their respective higher bodies.

TARGETS OF THE MANUAL

The manual is prepared to capacitate seed data collectors and seed data managers at different levels, namely, kebele, woreda, zone and region. Hence, the user of the manual must know the basics of using EXCEL.

TRAINING WORKSHOP PROGRAM

| Date | Time | Activity | Responsible person |
|------------------------------|-------------|--|--------------------------|
| Thursday Oct. 26, 2023 | Morning | Desktops transported to Assella | CRS Ethiopia |
| | 13:30-14:30 | Orientation on Desktops and program application | CRS |
| | 15:00-17:30 | Exercising with Excel Application | Dr. Bezabih Emana |
| Friday Oct.27, 2023 | 08:30-9:00 | Introduction and Objectives of the Training | Dr. Bezabih Emana |
| | 09:00-10:00 | Overview of the seed data collection instruments and key variables for analysis (Usual and Augmenting) | Dr. Bezabih+ Gimma |
| | 10:00-10:30 | Discussion | Participants |
| | 10:30-11:00 | Refreshment break | Organizers |
| | 11:00-12:30 | Data cleaning | Dr. Bezabih+ Gimma |
| | 12:30-13:00 | Data consistency | Dr. Bezabih+ Gimma |
| | 13:00-14:00 | Lunch break | Organizers |
| | 14:00-15:30 | Group exercise: data quality checking and detecting errors | Participants |
| | 15:30-16:00 | Refreshment break | Organizers |
| Saturday Oct. 28, 2023 | 16:00-17:00 | Data cleaning: Fixing errors | Dr. Bezabih+ Gimma |
| | 08:30-9:30 | Data aggregation | Dr. Bezabih+ Gimma |
| | 09:30-10:30 | Group exercise: data aggregation at different levels | Participants |
| | 10:30-11:00 | Refreshment break | Organizers |
| | 11:00-12:30 | Discussion on aggregation/summary making | Participants/facilitated |
| | 12:30-13:00 | Adjustment of the demand data | Dr. Bezabih+ Gimma |
| | 13:00-14:00 | Lunch break | Organizers |
| | 14:00-15:30 | Reporting data | Dr. Bezabih+ Gimma |
| | 15:30-16:00 | Refreshment break | Organizers |
| | 16:00-17:00 | Group exercise on data reporting at different levels | Participants |
| | 17:00-17:30 | Ways forward and concluding remark | Ebrahim + Bezabih |

OBJECTIVES OF THE TRAINING WORKSHOP

The aim of the workshop is creating awareness on the application of the seed demand data collection, data management and reporting and discussing on the ways forward.

This capacity building training enables the trainees to detect errors or inconsistencies in the dataset, clean the data and ensure data quality, analyze the data and report. The specific objectives are to enable the trainees to:

1. check consistencies and inconsistencies in the data
2. conduct data cleaning
3. ensure data quality
4. summarize data and
5. report the data

APPROACH

The dissemination activities covered presentation of the approach used to implement the seed demand data collection, the capacity building made and future application of the approach. This was done prior to the training on data management. The participants of the dissemination workshop were also trained on data management and evaluated the training as well as the outputs of the capacity building activities.

The training workshop was conducted in a participatory method:

- Power point presentation of the training manual was made by the consultants.
- Discussions on issues that need clarification were made.
- Group exercises were done using their own data. Under the supervision and follow up by the consultant, the group members exercised data management activities at different levels:
 - groups for kebele level data management
 - groups for woreda level data management
 - group for zone level data management
- Materials: Flip chart, LCD, note pads and pens were provided
- The outcomes of the group works were discussed in a plenary so that different groups learn from each other.

OVERVIEW OF THE SEED DEMAND DATA COLLECTION INSTRUMENTS

To inform the trainees what data needs to be collected and key variables for the collected data will be analyzed and reported. These include:

- **Identification Variables** – names of the zone, woreda, kebele, household (farmer), name of the cooperative (or group) for EGS data
- **Crops** for which seed is demanded: wheat, barley and faba bean
- **Usual variables:** aims to collect data on area for cultivation of the crops and quantity of seed
- **Augmenting variables:** own seed saved for planting – to be used to adjust the seed demand
- **Varieties:** Variety level disaggregated seed demand data

- **Sources of seed:** associated with seed quality
- **Quantity of output produced**
- **Price information**
- **More questions** were added to understand the seed demand situation in depth

Note: The data collection was done using Kobo Toolbox program. The simplified version of the questionnaire can be used for Excel Based Data Collection. That allows transporting the data directly from the Tablet to Desktop, for cleaning, analysis and reporting.

DATA CLEANING

Data cleaning is the process of ensuring data is correct, consistent, and usable. It is the process of detecting and correcting missing or inaccurate records in a dataset. It refers to identifying incomplete, incorrect, inaccurate, or irrelevant parts of the data and then replacing, modifying, or deleting the data.

Data cleaning is undertaken to ensure the quality of the data. It is conducted at different levels and for different purposes. After data is collected, data cleaning is conducted to reduce the errors that occurred during the data collection. It involves undertaking consistency checks and identifying missing values in the data.

Although data quality assurance activity begins before the commencement of data collection and continues during data collection, there are some measures undertaken to ensure the quality of the data after the data is collected.

Why is it important?

Data cleaning enables to ensure the consistency of the data, which is the basis for further use of the data. Data with high quality is the basis for better decision making and far-reaching impacts. In contrary, inconsistent data leads to wrong conclusions. Even limited errors and inconsistencies can make a whole data set useless for statistics. Decisions generated from inconsistent data will have bad consequences.

Incorrect or inconsistent data leads to false conclusions. And so, how well you clean and understand the data has a high impact on the quality of the results.

Therefore, having clean and consistent data will ultimately increase the quality decisions and its impacts. Data cleaning is an important activity in data management system.

COMMON ERRORS IN DATA

Data cleaning is about identifying the errors in the data set and addressing them. Data cleaning starts by identifying errors in the data.

Incomplete records: Due to haste during data collection and some negligence some fields could be jumped unfilled.

Wrong records: Filling inappropriate records. Numbers instead of text for text fields or vice versa.

Inappropriate records: exaggerations (both too high or too small values) due to poor probing during data collection or lack of proper attention. This could also be due to multiple unit of measurement (Example, quantity). In our data there are cases where records are made assuming quintals, instead of kgs.

Duplicates: Duplicates are data points that are repeated in your dataset.

It often happens when for example

- Data are combined from different sources
- The user may hit submit button twice thinking the form wasn't actually submitted.
- A request to online booking was submitted twice correcting wrong information that was entered accidentally in the first time.

A common symptom is when two users have the same identity number. Or, the same article was scrapped twice.

And therefore, they simply should be removed.

Type conversion: Make sure numbers are stored as numerical data types. A date should be stored as a date object, and so on.

Pad strings: Strings can be padded with spaces or other characters to a certain width. For example, some numerical codes are often represented with prepending zeros to ensure they always have the same number of digits.

Fix typos: Strings can be entered in many different ways, and no wonder, can have mistakes

Categorical values viz continuous values: Categorical values can be converted into and from numbers if needed.

Outliers: They are values that are significantly different from all other observations. Any data value that lies more than $(1.5 * IQR)$ away from the Q1 and Q3 quartiles is considered an outlier. Outliers should not be removed unless there is a good reason for that.

Quartiles segment any distribution that's ordered from low to high into four equal parts. The interquartile range (IQR) contains the second and third quartiles, or the middle half of your data set.

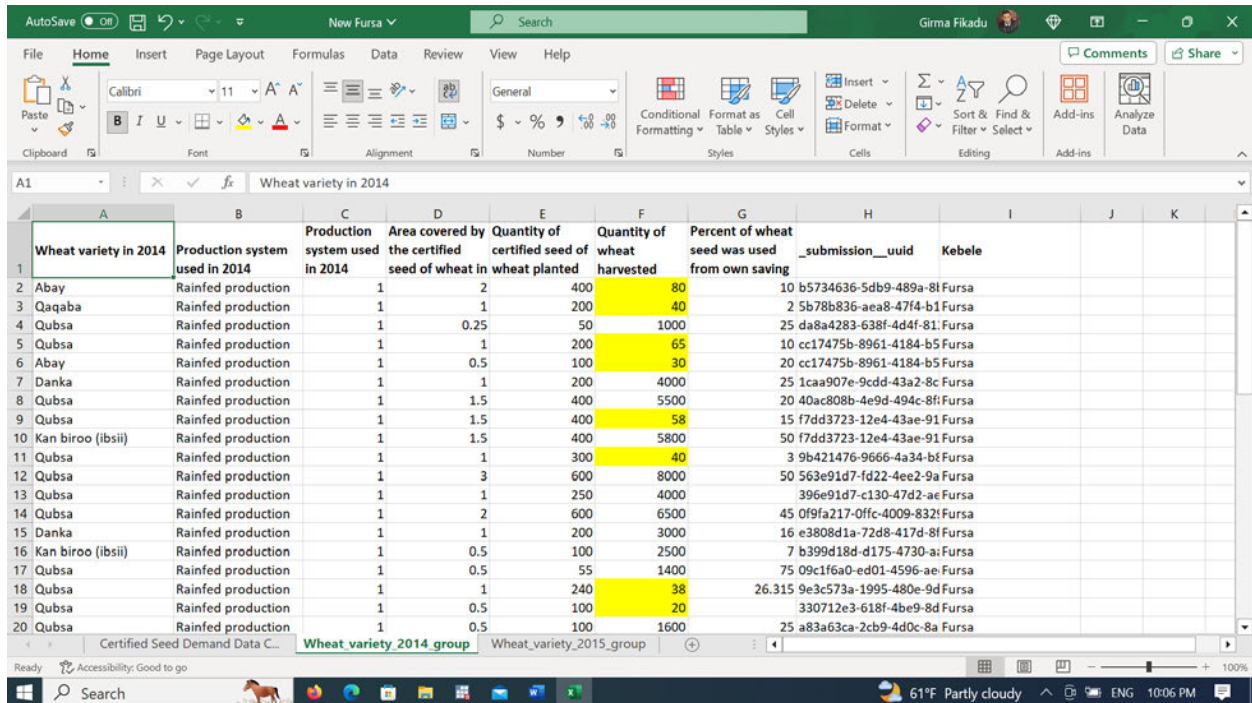
For example, one can notice some weird, suspicious values that are unlikely to happen, and so decides to remove them.

IDENTIFYING ERRORS

As mentioned above, data cleaning involves identifying errors in data. Missing values, exaggerations and mismatches in the data need to be properly identified. Some of the ways by which errors in data can be identified and cleaned include:

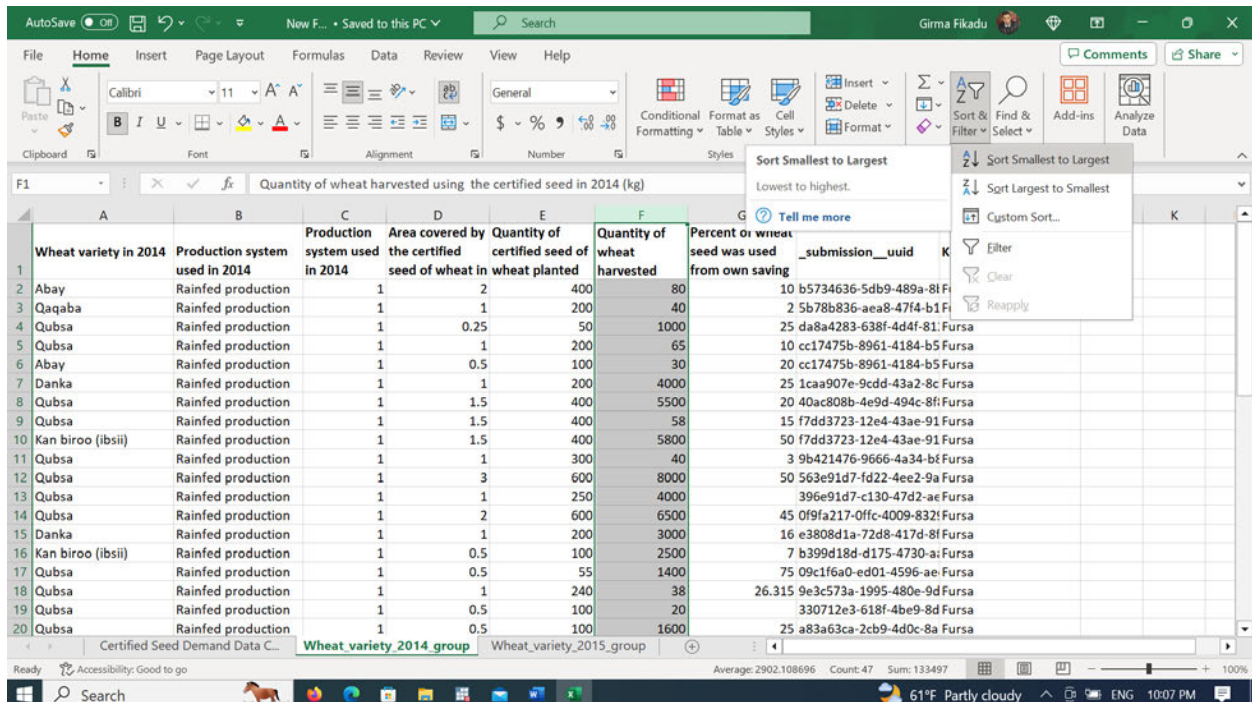
Visualize your data: Some of the errors in the data can be identified by going through the data and visually observing them. This activity is particularly feasible when the data set is small. In the following example, one can easily identify the missing value by visually observing the data.

One can identify by observing the data that the figures highlighted with yellow have errors. In this case, the amount harvested is lower than the amount of seed used.



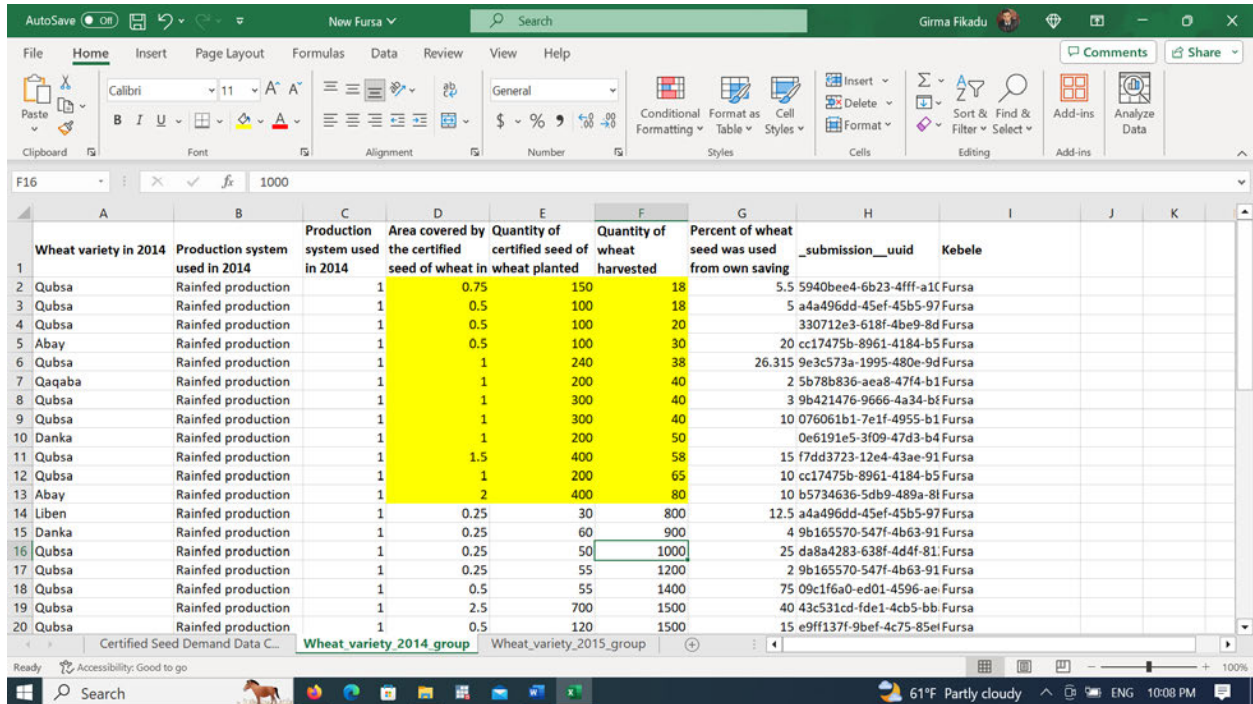
| Wheat variety in 2014 | Production system used in 2014 | Production system used in 2014 | Area covered by the certified seed of wheat in | Quantity of certified seed of wheat planted | Quantity of wheat harvested | Percent of wheat seed was used from own saving | _submission_uid | Kebele |
|-----------------------|--------------------------------|--------------------------------|--|---|-----------------------------|--|------------------------|--------|
| Abay | Rainfed production | 1 | 2 | 400 | 80 | 10 | b5734636-5db9-489a-8f | Fursa |
| Qaqaba | Rainfed production | 1 | 1 | 200 | 40 | 2 | 5b78b836-aea8-47f4-b1 | Fursa |
| Qubsa | Rainfed production | 1 | 0.25 | 50 | 1000 | 25 | da8a4283-638f-4d4f-81 | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 200 | 65 | 10 | cc17475b-8961-4184-b5 | Fursa |
| Abay | Rainfed production | 1 | 0.5 | 100 | 30 | 20 | cc17475b-8961-4184-b5 | Fursa |
| Danka | Rainfed production | 1 | 1 | 200 | 4000 | 25 | 1caa907e-9cdd-43a2-8c | Fursa |
| Qubsa | Rainfed production | 1 | 1.5 | 400 | 5500 | 20 | 40ac808b-4e9d-494c-8f | Fursa |
| Qubsa | Rainfed production | 1 | 1.5 | 400 | 58 | 15 | f7dd3723-12e4-43ae-91 | Fursa |
| Kan biroo (ibsiil) | Rainfed production | 1 | 1.5 | 400 | 5800 | 50 | f7dd3723-12e4-43ae-91 | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 300 | 40 | 3 | 9b421476-9666-4a34-bf | Fursa |
| Qubsa | Rainfed production | 1 | 3 | 600 | 8000 | 50 | 563e91d7-fd22-4ee2-9a | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 250 | 4000 | 396e91d7-c130-47d2-ae | Fursa | |
| Qubsa | Rainfed production | 1 | 2 | 600 | 6500 | 45 | 0f9fa217-0ffc-4009-832 | Fursa |
| Danka | Rainfed production | 1 | 1 | 200 | 3000 | 16 | e3808d1a-72d8-417d-8f | Fursa |
| Kan biroo (ibsiil) | Rainfed production | 1 | 0.5 | 100 | 2500 | 7 | b399d18d-d175-4730-a | Fursa |
| Qubsa | Rainfed production | 1 | 0.5 | 55 | 1400 | 75 | 09c1f6a0-ed01-4596-ae | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 240 | 38 | 26.315 | 9e3c573a-1995-480e-9d | Fursa |
| Qubsa | Rainfed production | 1 | 0.5 | 100 | 20 | 330712e3-618f-4be9-8d | Fursa | |
| Qubsa | Rainfed production | 1 | 0.5 | 100 | 1600 | 25 | a83a63ca-2cb9-4d0c-8a | Fursa |

Filter: By going through the variables in the data and conducting filtering and sorting some inconsistencies and in appropriate data values can be identified. Such activity can identify exaggerations (more than usual data)



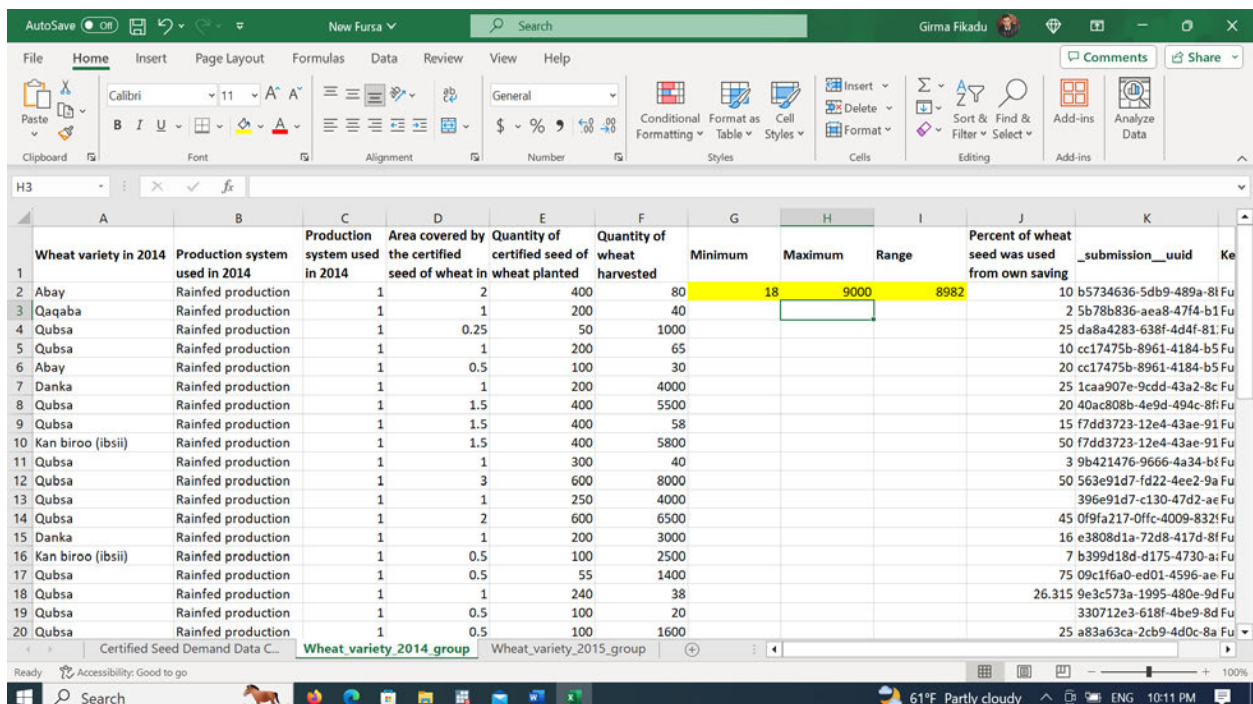
| Wheat variety in 2014 | Production system used in 2014 | Production system used in 2014 | Area covered by the certified seed of wheat in | Quantity of certified seed of wheat planted | Quantity of wheat harvested | Percent of wheat seed was used from own saving | _submission_uid | Kebele |
|-----------------------|--------------------------------|--------------------------------|--|---|-----------------------------|--|------------------------|--------|
| Qubsa | Rainfed production | 1 | 0.5 | 100 | 30 | 20 | cc17475b-8961-4184-b5 | Fursa |
| Qubsa | Rainfed production | 1 | 0.5 | 100 | 20 | 330712e3-618f-4be9-8d | Fursa | |
| Qubsa | Rainfed production | 1 | 0.5 | 100 | 1600 | 25 | a83a63ca-2cb9-4d0c-8a | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 240 | 38 | 26.315 | 9e3c573a-1995-480e-9d | Fursa |
| Qubsa | Rainfed production | 1 | 0.5 | 55 | 1400 | 75 | 09c1f6a0-ed01-4596-ae | Fursa |
| Kan biroo (ibsiil) | Rainfed production | 1 | 0.5 | 100 | 2500 | 7 | b399d18d-d175-4730-a | Fursa |
| Danka | Rainfed production | 1 | 1 | 200 | 3000 | 16 | e3808d1a-72d8-417d-8f | Fursa |
| Qubsa | Rainfed production | 1 | 2 | 600 | 6500 | 45 | 0f9fa217-0ffc-4009-832 | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 250 | 4000 | 396e91d7-c130-47d2-ae | Fursa | |
| Qubsa | Rainfed production | 1 | 3 | 600 | 8000 | 50 | 563e91d7-fd22-4ee2-9a | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 300 | 40 | 3 | 9b421476-9666-4a34-bf | Fursa |
| Kan biroo (ibsiil) | Rainfed production | 1 | 1.5 | 400 | 5800 | 50 | f7dd3723-12e4-43ae-91 | Fursa |
| Qubsa | Rainfed production | 1 | 1.5 | 400 | 58 | 15 | f7dd3723-12e4-43ae-91 | Fursa |
| Qubsa | Rainfed production | 1 | 1.5 | 400 | 5500 | 20 | 40ac808b-4e9d-494c-8f | Fursa |
| Danka | Rainfed production | 1 | 1 | 200 | 4000 | 25 | 1caa907e-9cdd-43a2-8c | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 200 | 65 | 10 | cc17475b-8961-4184-b5 | Fursa |
| Qubsa | Rainfed production | 1 | 0.25 | 50 | 1000 | 25 | da8a4283-638f-4d4f-81 | Fursa |
| Qaqaba | Rainfed production | 1 | 1 | 200 | 40 | 2 | 5b78b836-aea8-47f4-b1 | Fursa |
| Abay | Rainfed production | 1 | 2 | 400 | 80 | 10 | b5734636-5db9-489a-8f | Fursa |

The sorted data looks the following. It is sorted from the smallest to the largest. Upon comparing the three data inputs in the three variables (Area, seed amount, and amount harvested), one can recognize that the figures under amount harvested have errors.



| Wheat variety in 2014 | Production system used in 2014 | Production system used in 2014 | Area covered by the certified seed of wheat in | Quantity of certified seed of wheat planted | Quantity of wheat harvested | Percent of wheat seed was used from own saving | _submission_uid | Kebele |
|-----------------------|--------------------------------|--------------------------------|--|---|-----------------------------|--|-----------------------|-------------------------|
| 1 | Qubsa | Rainfed production | 1 | 0.75 | 150 | 18 | 5.5 | 5940bee4-6b23-4fff-a1c |
| 2 | Qubsa | Rainfed production | 1 | 0.5 | 100 | 18 | 5 | a4a496dd-45ef-45b5-97 |
| 3 | Qubsa | Rainfed production | 1 | 0.5 | 100 | 20 | 330712e3-618f-4be9-8d | |
| 4 | Abay | Rainfed production | 1 | 0.5 | 100 | 30 | 20 | cc17475b-8961-4184-b5 |
| 5 | Qubsa | Rainfed production | 1 | 1 | 240 | 38 | 26.315 | 9e3c573a-1995-480e-9d |
| 6 | Qaqaba | Rainfed production | 1 | 1 | 200 | 40 | 2 | 5b78b836-aea8-47f4-b1 |
| 7 | Qubsa | Rainfed production | 1 | 1 | 300 | 40 | 3 | 9b421476-9666-4a34-bf |
| 8 | Qubsa | Rainfed production | 1 | 1 | 300 | 40 | 10 | 076061b1-7e1f-4955-b1 |
| 9 | Danka | Rainfed production | 1 | 1 | 200 | 50 | 0e6191e5-3f09-47d3-b4 | |
| 10 | Qubsa | Rainfed production | 1 | 1.5 | 400 | 58 | 15 | f7dd3723-12e4-43ae-91 |
| 11 | Qubsa | Rainfed production | 1 | 1 | 200 | 65 | 10 | cc17475b-8961-4184-b5 |
| 12 | Abay | Rainfed production | 1 | 2 | 400 | 80 | 10 | b5734636-5db9-489a-8f |
| 13 | Liben | Rainfed production | 1 | 0.25 | 30 | 800 | 12.5 | a4a496dd-45ef-45b5-97 |
| 14 | Danka | Rainfed production | 1 | 0.25 | 60 | 900 | 4 | 9b165570-547f-4b63-91 |
| 15 | Qubsa | Rainfed production | 1 | 0.25 | 50 | 1000 | 25 | da8a4283-638f-4d4f-81 |
| 16 | Qubsa | Rainfed production | 1 | 0.25 | 55 | 1200 | 2 | 9b165570-547f-4b63-91 |
| 17 | Qubsa | Rainfed production | 1 | 0.5 | 55 | 1400 | 75 | 09c1f6a0-ed01-4596-ae |
| 18 | Qubsa | Rainfed production | 1 | 2.5 | 700 | 1500 | 40 | 43c531cd-fde1-4cb5-bb |
| 19 | Qubsa | Rainfed production | 1 | 0.5 | 120 | 1500 | 15 | e9ff137f-9bef-4c75-85ef |
| 20 | Qubsa | Rainfed production | 1 | 0.5 | 120 | 1500 | 15 | e9ff137f-9bef-4c75-85ef |

Run summary tables: Running summary tables such as frequencies can show missing values in the data. Running range values can also show the extremes in the data. For example, upon running ranges, one can notice that the range is high, meaning that there is huge difference between the highest and lowest value.



| Wheat variety in 2014 | Production system used in 2014 | Production system used in 2014 | Area covered by the certified seed of wheat in | Quantity of certified seed of wheat planted | Quantity of wheat harvested | Minimum | Maximum | Range | Percent of wheat seed was used from own saving | _submission_uid | Ke |
|-----------------------|--------------------------------|--------------------------------|--|---|-----------------------------|---------|---------|-------|--|-----------------|-------------------------|
| 1 | Abay | Rainfed production | 1 | 2 | 400 | 80 | 18 | 9000 | 8982 | 10 | b5734636-5db9-489a-8f |
| 2 | Qaqaba | Rainfed production | 1 | 1 | 200 | 40 | | | | 2 | 5b78b836-aea8-47f4-b1 |
| 3 | Qubsa | Rainfed production | 1 | 0.25 | 50 | 1000 | | | | 25 | da8a4283-638f-4d4f-81 |
| 4 | Qubsa | Rainfed production | 1 | 1 | 200 | 65 | | | | 20 | cc17475b-8961-4184-b5 |
| 5 | Abay | Rainfed production | 1 | 0.5 | 100 | 30 | | | | 20 | cc17475b-8961-4184-b5 |
| 6 | Danka | Rainfed production | 1 | 1 | 200 | 4000 | | | | 25 | 1caa907e-9cdd-43a2-8c |
| 7 | Qubsa | Rainfed production | 1 | 1.5 | 400 | 5500 | | | | 20 | 40ac808b-4e9d-494c-8f |
| 8 | Qubsa | Rainfed production | 1 | 1.5 | 400 | 58 | | | | 15 | f7dd3723-12e4-43ae-91 |
| 9 | Kan biroo (ibsi) | Rainfed production | 1 | 1.5 | 400 | 5800 | | | | 50 | f7dd3723-12e4-43ae-91 |
| 10 | Qubsa | Rainfed production | 1 | 1 | 300 | 40 | | | | 3 | 9b421476-9666-4a34-bf |
| 11 | Qubsa | Rainfed production | 1 | 3 | 600 | 8000 | | | | 50 | 563e91d7-fd22-4ee2-9a |
| 12 | Qubsa | Rainfed production | 1 | 1 | 250 | 4000 | | | | | 396e91d7-c130-47d2-ae |
| 13 | Qubsa | Rainfed production | 1 | 2 | 600 | 6500 | | | | 45 | 0f9fa217-0ffc-4009-832f |
| 14 | Danka | Rainfed production | 1 | 1 | 200 | 3000 | | | | 16 | e3808d1a-72d8-417d-8f |
| 15 | Kan biroo (ibsi) | Rainfed production | 1 | 0.5 | 100 | 2500 | | | | 7 | b399d18d-d175-4730-a |
| 16 | Qubsa | Rainfed production | 1 | 0.5 | 55 | 1400 | | | | 75 | 09c1f6a0-ed01-4596-ae |
| 17 | Qubsa | Rainfed production | 1 | 1 | 240 | 38 | | | | 26.315 | 9e3c573a-1995-480e-9d |
| 18 | Qubsa | Rainfed production | 1 | 0.5 | 100 | 20 | | | | | 330712e3-618f-4be9-8d |
| 19 | Qubsa | Rainfed production | 1 | 0.5 | 100 | 1600 | | | | 25 | a83a63ca-2cb9-4d0c-8a |
| 20 | Qubsa | Rainfed production | 1 | 0.5 | 100 | 1600 | | | | 25 | a83a63ca-2cb9-4d0c-8a |

Group Exercise #1: Form 3 groups each to identify errors in seed data collected in Arsi

MEASURES TO CLEAN DATA

After errors in the data are identified, the next step is about deciding how to deal with those errors. The following measures could be taken based on the cause of the error.

Missing values

Given the fact the missing values are unavoidable leaves us with the question of what to do when we encounter them. Ignoring the missing data is the same as digging holes in a boat; It will sink.

There are three, or perhaps more, ways to deal with them.

- **Option 1: Drop.**

If the missing values in a column rarely happen and occur at random, then the easiest and most forward solution is to drop observations (rows) that have missing values.

If most of the column's values are missing, and occur at random, then a typical decision is to drop the whole column.

This is particularly useful when doing statistical analysis, since filling in the missing values may yield unexpected or biased results.

- **Two. Impute.**

It means to calculate the missing value based on other observations. There are quite a lot of methods to do that. **First** one is using **statistical values** like mean, median. However, none of these guarantees unbiased data, especially if there are many missing values.

Mean is most useful when the original data is **not skewed**, while the **median is more robust**, not sensitive to outliers, and thus used when **data is skewed**.

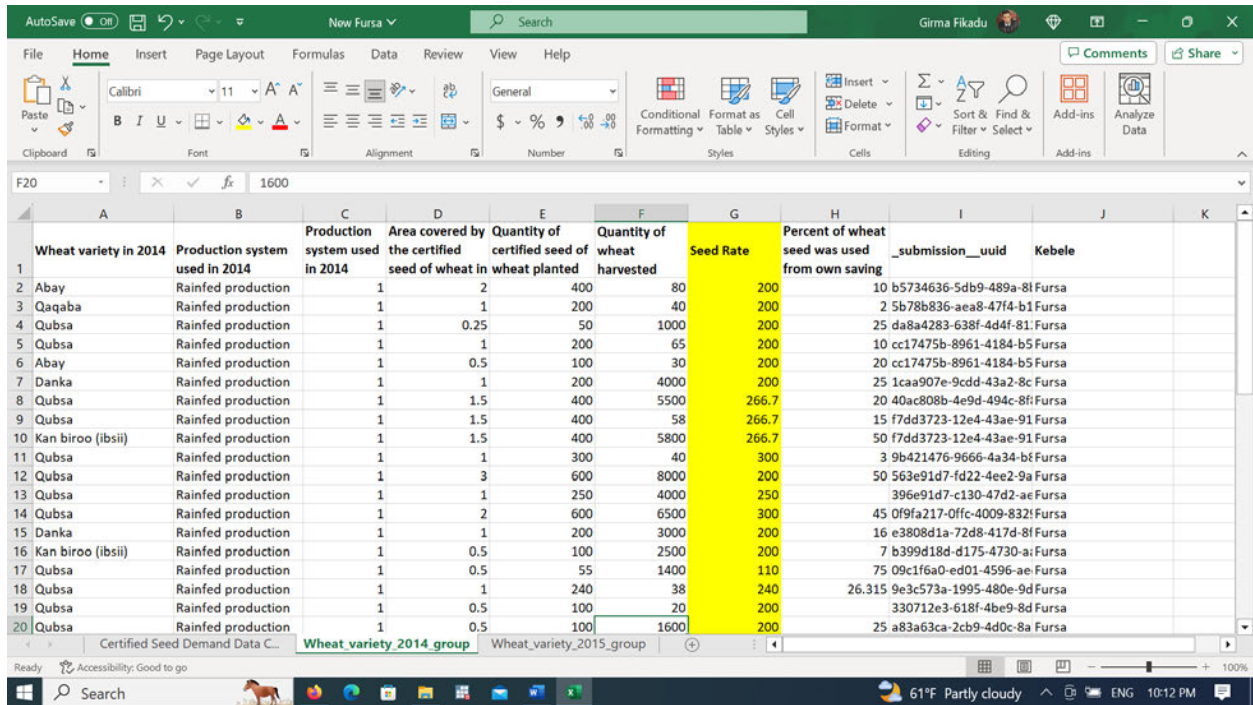
- **Third. Hot-deck.**

Copying values from other similar records. This is only useful if you have enough available data. And, it can be applied to numerical and categorical data.

Correcting the record: Some of the errors identified can easily be corrected by identifying the cause of the error. In the above examples, the error of smaller values under amount of harvest is attributed to the unit of measurement. Some of the records are made considering quintal (100kg). Therefore, seed for additional information such as contacting the person responsible for data collection, the figures will be converted to kilograms instead of quintals.

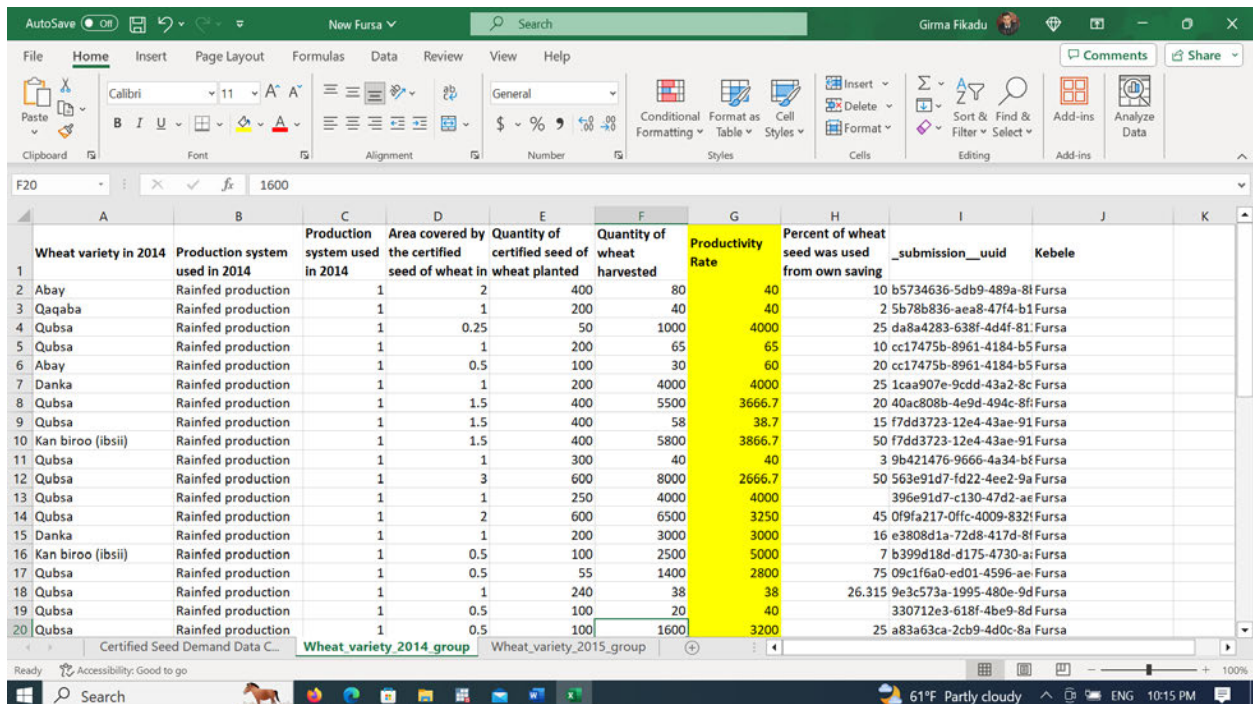
Moreover, correction could also be done by triangulating it with other sources. By comparing them with other sources, such as trends in previous years, corrections could be made. For the data collected in Arsi zone, the seed rate could be calculated and compared with standards as well as previous year data.

The seed rate calculation shows a range of 110 to 300 kg of seed applied per 1 hectare of land.



| Wheat variety in 2014 | Production system used in 2014 | Production system used in 2014 | Area covered by the certified seed of wheat in | Quantity of certified seed of wheat planted | Quantity of wheat harvested | Seed Rate | Percent of wheat seed was used from own saving | _submission_uuid | Kebele |
|-----------------------|--------------------------------|--------------------------------|--|---|-----------------------------|-----------|--|------------------------------------|--------|
| 1 | Abay | Rainfed production | 1 | 2 | 400 | 80 | 200 | 10 b5734636-5db9-489a-8f Fursa | |
| 2 | Qaqaba | Rainfed production | 1 | 1 | 200 | 40 | 200 | 2 5b78b836-aea8-47f4-b1 Fursa | |
| 3 | Qbsa | Rainfed production | 1 | 0.25 | 50 | 1000 | 200 | 25 da8a4283-638f-4d4f-81 Fursa | |
| 4 | Qbsa | Rainfed production | 1 | 1 | 200 | 65 | 200 | 10 cc17475b-8961-4184-b5 Fursa | |
| 5 | Abay | Rainfed production | 1 | 0.5 | 100 | 30 | 200 | 20 cc17475b-8961-4184-b5 Fursa | |
| 6 | Danka | Rainfed production | 1 | 1 | 200 | 4000 | 200 | 25 1caa907e-9cdd-43a2-8c Fursa | |
| 7 | Qbsa | Rainfed production | 1 | 1.5 | 400 | 5500 | 266.7 | 20 40ac808b-4e9d-494c-8f Fursa | |
| 8 | Qbsa | Rainfed production | 1 | 1.5 | 400 | 58 | 266.7 | 15 f7dd3723-12e4-43ae-91 Fursa | |
| 9 | Kan biroo (ibsiil) | Rainfed production | 1 | 1.5 | 400 | 5800 | 266.7 | 50 f7dd3723-12e4-43ae-91 Fursa | |
| 10 | Qbsa | Rainfed production | 1 | 1 | 300 | 40 | 300 | 3 9b421476-9666-4a34-bf Fursa | |
| 11 | Qbsa | Rainfed production | 1 | 3 | 600 | 8000 | 200 | 50 563e91d7-fd22-4ee2-9a Fursa | |
| 12 | Qbsa | Rainfed production | 1 | 1 | 250 | 4000 | 250 | 396e91d7-c130-47d2-ae Fursa | |
| 13 | Qbsa | Rainfed production | 1 | 2 | 600 | 6500 | 300 | 45 0f9fa217-0ffc-4009-832f Fursa | |
| 14 | Danka | Rainfed production | 1 | 1 | 200 | 3000 | 200 | 16 e3808d1a-72d8-417d-8f Fursa | |
| 15 | Kan biroo (ibsiil) | Rainfed production | 1 | 0.5 | 100 | 2500 | 200 | 7 b399d18d-d175-4730-af Fursa | |
| 16 | Qbsa | Rainfed production | 1 | 0.5 | 55 | 1400 | 110 | 75 09c1f6a0-ed01-4596-ae Fursa | |
| 17 | Qbsa | Rainfed production | 1 | 1 | 240 | 38 | 240 | 26.315 9e3c573a-1995-480e-9d Fursa | |
| 18 | Qbsa | Rainfed production | 1 | 0.5 | 100 | 20 | 200 | 330712e3-618f-4be9-8d Fursa | |
| 19 | Qbsa | Rainfed production | 1 | 0.5 | 100 | 1600 | 200 | 25 a83a63ca-2cb9-4d0c-8a Fursa | |
| 20 | Qbsa | Rainfed production | 1 | 0.5 | 100 | 1600 | 200 | 25 a83a63ca-2cb9-4d0c-8a Fursa | |

Similarly, you can calculate yield (output per ha). Too much variation in yield may indicate errors in the original data.



| Wheat variety in 2014 | Production system used in 2014 | Production system used in 2014 | Area covered by the certified seed of wheat in | Quantity of certified seed of wheat planted | Quantity of wheat harvested | Productivity Rate | Percent of wheat seed was used from own saving | _submission_uuid | Kebele |
|-----------------------|--------------------------------|--------------------------------|--|---|-----------------------------|-------------------|--|------------------------------------|--------|
| 1 | Abay | Rainfed production | 1 | 2 | 400 | 80 | 40 | 10 b5734636-5db9-489a-8f Fursa | |
| 2 | Qaqaba | Rainfed production | 1 | 1 | 200 | 40 | 40 | 2 5b78b836-aea8-47f4-b1 Fursa | |
| 3 | Qbsa | Rainfed production | 1 | 0.25 | 50 | 1000 | 4000 | 25 da8a4283-638f-4d4f-81 Fursa | |
| 4 | Qbsa | Rainfed production | 1 | 1 | 200 | 65 | 65 | 10 cc17475b-8961-4184-b5 Fursa | |
| 5 | Abay | Rainfed production | 1 | 0.5 | 100 | 30 | 60 | 20 cc17475b-8961-4184-b5 Fursa | |
| 6 | Danka | Rainfed production | 1 | 1 | 200 | 4000 | 4000 | 25 1caa907e-9cdd-43a2-8c Fursa | |
| 7 | Qbsa | Rainfed production | 1 | 1.5 | 400 | 5500 | 3666.7 | 20 40ac808b-4e9d-494c-8f Fursa | |
| 8 | Qbsa | Rainfed production | 1 | 1.5 | 400 | 58 | 38.7 | 15 f7dd3723-12e4-43ae-91 Fursa | |
| 9 | Kan biroo (ibsiil) | Rainfed production | 1 | 1.5 | 400 | 5800 | 3866.7 | 50 f7dd3723-12e4-43ae-91 Fursa | |
| 10 | Qbsa | Rainfed production | 1 | 1 | 300 | 40 | 40 | 3 9b421476-9666-4a34-bf Fursa | |
| 11 | Qbsa | Rainfed production | 1 | 3 | 600 | 8000 | 2666.7 | 50 563e91d7-fd22-4ee2-9a Fursa | |
| 12 | Qbsa | Rainfed production | 1 | 1 | 250 | 4000 | 4000 | 396e91d7-c130-47d2-ae Fursa | |
| 13 | Qbsa | Rainfed production | 1 | 2 | 600 | 6500 | 3250 | 45 0f9fa217-0ffc-4009-832f Fursa | |
| 14 | Danka | Rainfed production | 1 | 1 | 200 | 3000 | 3000 | 16 e3808d1a-72d8-417d-8f Fursa | |
| 15 | Kan biroo (ibsiil) | Rainfed production | 1 | 0.5 | 100 | 2500 | 5000 | 7 b399d18d-d175-4730-af Fursa | |
| 16 | Qbsa | Rainfed production | 1 | 0.5 | 55 | 1400 | 2800 | 75 09c1f6a0-ed01-4596-ae Fursa | |
| 17 | Qbsa | Rainfed production | 1 | 1 | 240 | 38 | 38 | 26.315 9e3c573a-1995-480e-9d Fursa | |
| 18 | Qbsa | Rainfed production | 1 | 0.5 | 100 | 20 | 40 | 330712e3-618f-4be9-8d Fursa | |
| 19 | Qbsa | Rainfed production | 1 | 0.5 | 100 | 1600 | 3200 | 25 a83a63ca-2cb9-4d0c-8a Fursa | |
| 20 | Qbsa | Rainfed production | 1 | 0.5 | 100 | 1600 | 3200 | 25 a83a63ca-2cb9-4d0c-8a Fursa | |

DATA QUALITY ASSURANCE

Verification helps assuring data quality. When done with the data cleaning, one should verify correctness by re-inspecting the data and making sure it rules and constraints do hold. After filling out the missing data, they might violate any of the rules and constraints. It might involve some manual correction if not possible otherwise.

Group Exercise #2: Form 3 groups each to take actions to improve data quality using seed data collected in Arsi

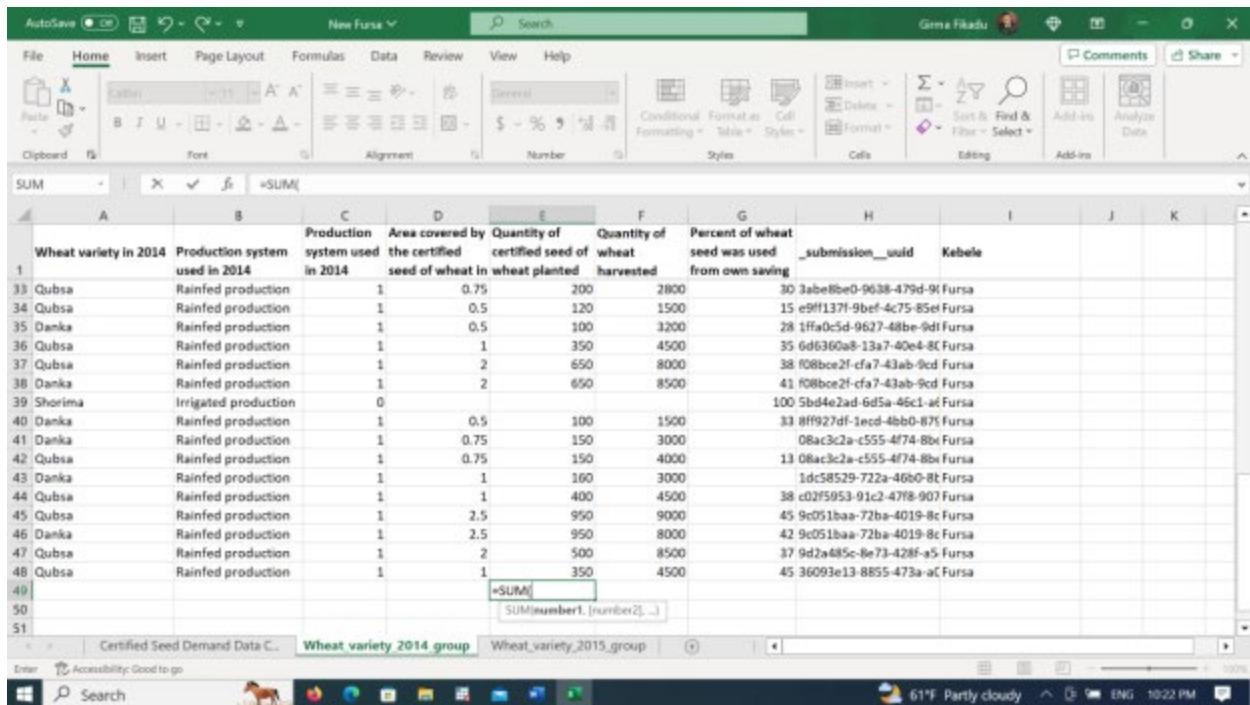
DATA AGGREGATION/SUMMARY

Once the data is cleaned. Data summary is done at different levels: Kebele, woreda, and zone levels.

Calculating totals: Total needs to be calculated for area, seed amount and harvested amount by summing all the values under each variable. Real time data collected from one of the kebeles in Arsi zone, Lode Hetosa kebele is used to demonstrate the aggregation. In the same way, aggregation of data shall be done at woreda and zone level.

The following are the steps:

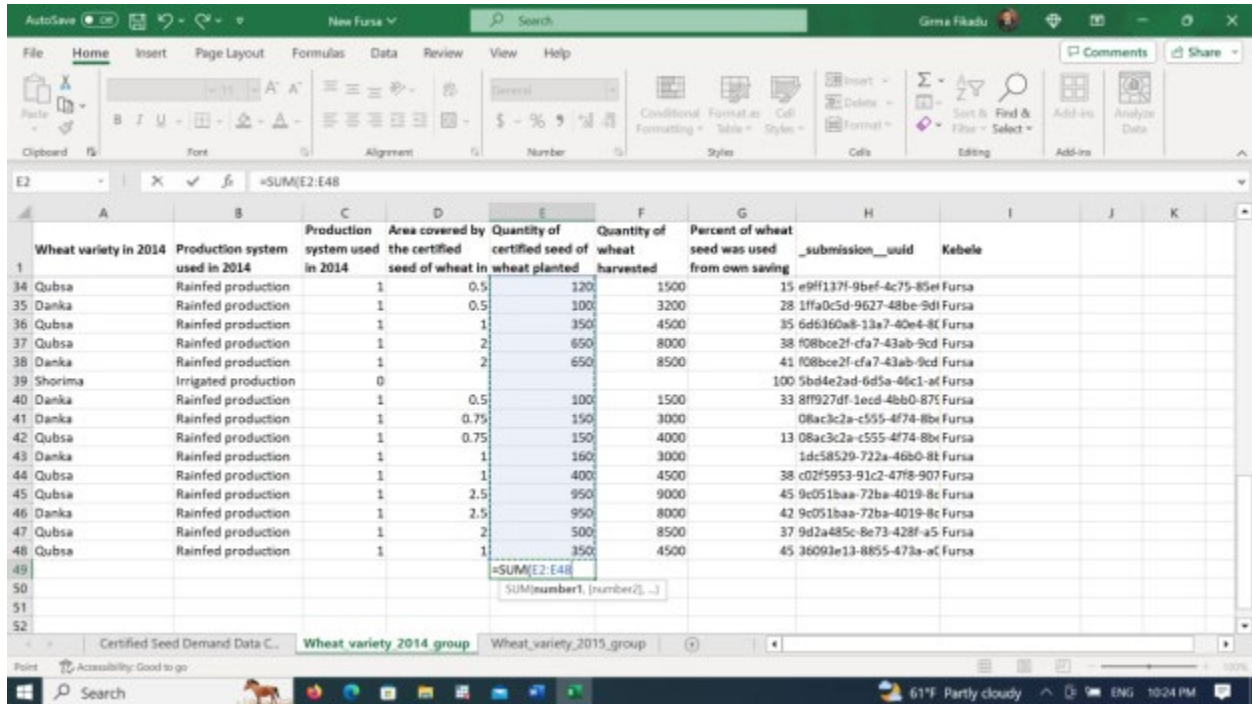
1st: Go to the cell at the end of the values under each variable (in this case, seed amount used) and write the following function: =SUM



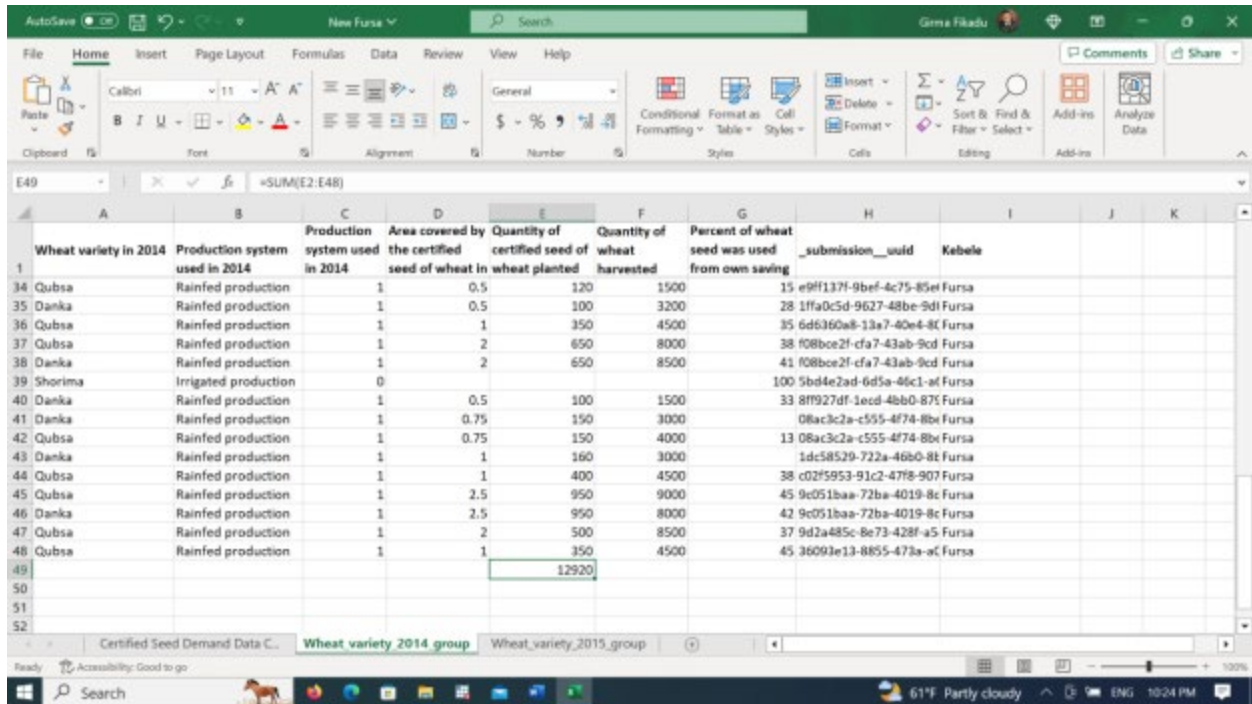
| | A | B | C | D | E | F | G | H | I | J | K |
|----|-----------------------|--------------------------------|--------------------------------|--|---|-----------------------------|--|------------------------|--------|---|---|
| | Wheat variety in 2014 | Production system used in 2014 | Production system used in 2014 | Area covered by the certified seed of wheat in | Quantity of certified seed of wheat planted | Quantity of wheat harvested | Percent of wheat seed was used from own saving | _submission_uuid | Kebele | | |
| 33 | Qubsa | Rainfed production | 1 | 0.75 | 200 | 2800 | 30 | 3abe8be0-9638-479d-9f | Fursa | | |
| 34 | Qubsa | Rainfed production | 1 | 0.5 | 120 | 1500 | 15 | e9ff137f-9bef-4c75-85e | Fursa | | |
| 35 | Danka | Rainfed production | 1 | 0.5 | 100 | 3200 | 28 | 1ffa0c5d-9627-48be-9d | Fursa | | |
| 36 | Qubsa | Rainfed production | 1 | 1 | 350 | 4500 | 35 | 6d6360a8-13a7-40e4-8c | Fursa | | |
| 37 | Qubsa | Rainfed production | 1 | 2 | 650 | 8000 | 38 | f08bce2f-cfa7-43ab-9cd | Fursa | | |
| 38 | Danka | Rainfed production | 1 | 2 | 650 | 8500 | 41 | f08bce2f-cfa7-43ab-9cd | Fursa | | |
| 39 | Shorima | Irrigated production | 0 | | | | 100 | 5bd4e2ad-6d5a-46c1-a | Fursa | | |
| 40 | Danka | Rainfed production | 1 | 0.5 | 100 | 1500 | 33 | 8ff927df-1ecd-4bb0-87 | Fursa | | |
| 41 | Danka | Rainfed production | 1 | 0.75 | 150 | 3000 | 08ac3c2a-c555-4f74-8b | Fursa | | | |
| 42 | Qubsa | Rainfed production | 1 | 0.75 | 150 | 4000 | 13 | 08ac3c2a-c555-4f74-8b | Fursa | | |
| 43 | Danka | Rainfed production | 1 | 1 | 100 | 3000 | 1dc58529-722a-46b0-8f | Fursa | | | |
| 44 | Qubsa | Rainfed production | 1 | 1 | 400 | 4500 | 38 | c02f5953-91c2-47f8-90 | Fursa | | |
| 45 | Qubsa | Rainfed production | 1 | 2.5 | 950 | 9000 | 45 | 9c051baa-72ba-4019-8c | Fursa | | |
| 46 | Danka | Rainfed production | 1 | 2.5 | 950 | 8000 | 42 | 9c051baa-72ba-4019-8c | Fursa | | |
| 47 | Qubsa | Rainfed production | 1 | 2 | 500 | 8500 | 37 | 9d2a485c-8e73-428f-a5 | Fursa | | |
| 48 | Qubsa | Rainfed production | 1 | 1 | 350 | 4500 | 45 | 36093e13-8855-473a-a | Fursa | | |
| 49 | | | | | =SUM | | | | | | |
| 50 | | | | | SUM(number1, [number2], ...) | | | | | | |
| 51 | | | | | | | | | | | |

2nd: Write the sum function and select the data ranges under the variable. It looks the following:

=SUM(E2:E48)



3rd After running the above function results in the following figure.



The figure shows that 12,920 kg of wheat seed is needed by the interviewed households, who are about 47 in number.

During the data collection, the seed demand data is collected from portion of wheat producing farmers in the area.

Therefore, to arrive at the amount of wheat seed demanded at the kebele level (for all of wheat producing farmers in the kebele), we need to calculate the amount of wheat seed demand at individual farmer level for the sampled households.

To arrive at the total wheat seed demand for the total wheat farmers in the kebele, multiply the individual figure for amount of wheat seed demand by the total number of wheat producing farmers in the kebele.

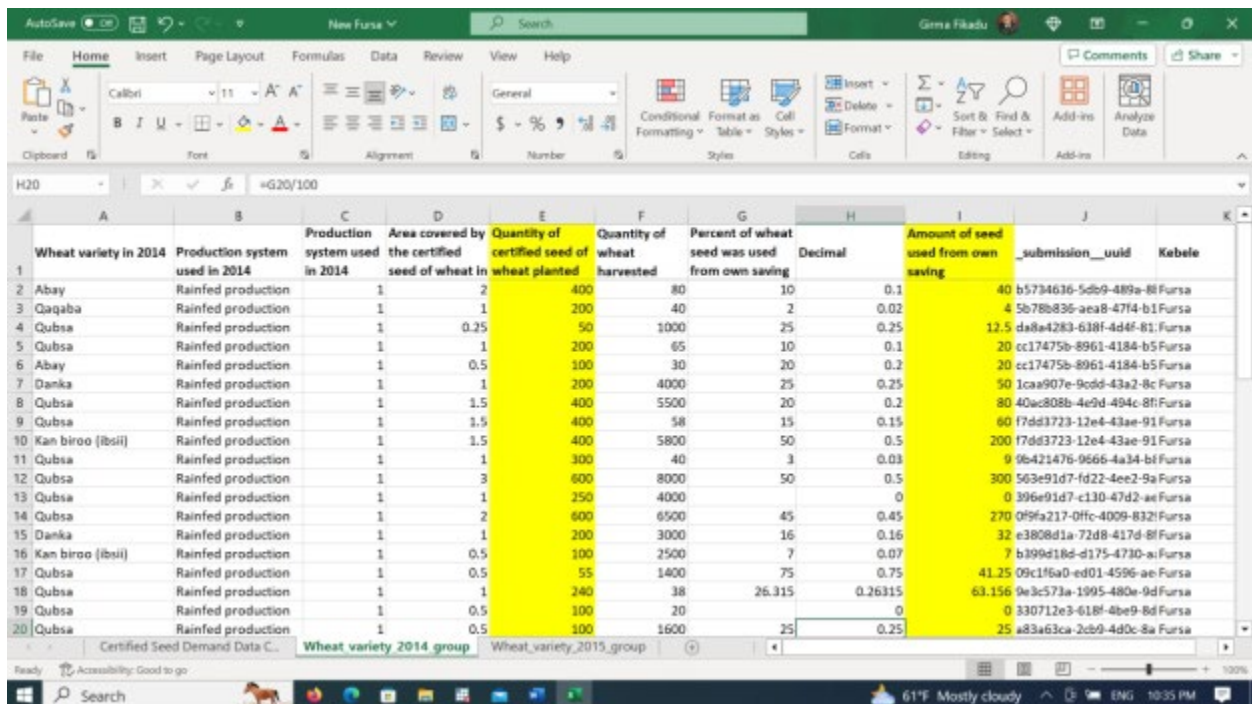
Average amount of wheat seed demand= 274.9 kg (=12,920kg/47)

Secondary data collected from Fursa kebele indicate that there are 283 wheat producing farmers in the kebele. Considering this figure, the total wheat seed demand for the kebele will be, 77,796.7 kg (=274.9*283).

AUGMENTED DATA:

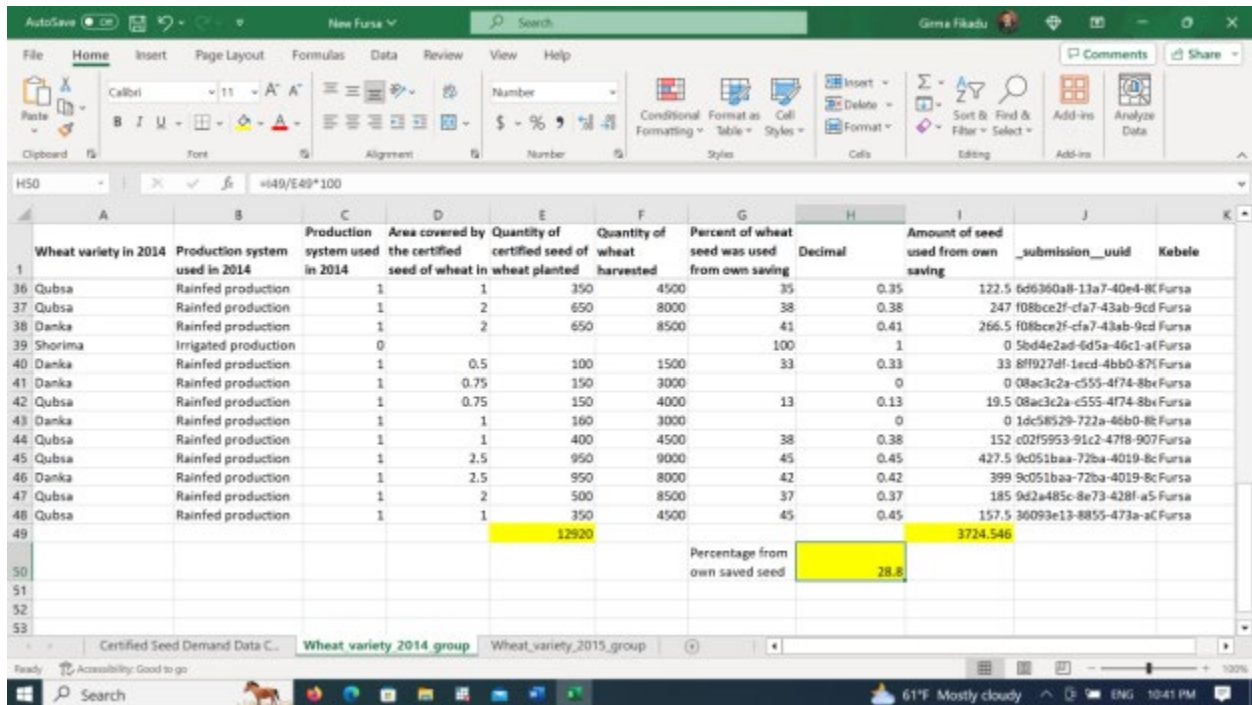
It is worth noting here that part of the seed will be covered from own saving. Therefore, to arrive at the net seed demand, the part that will be covered from own saving should be deducted from the total amount indicated above.

The highlighted area below shows amount of seed used in comparison with the amount covered from own saving.



| Wheat variety in 2014 | Production system used in 2014 | Production system used in 2014 | Area covered by the certified seed of wheat in | Quantity of certified seed of wheat planted | Quantity of wheat harvested | Percent of wheat seed was used from own saving | Decimal | Amount of seed used from own saving | _submission_uid | Kebele |
|-----------------------|--------------------------------|--------------------------------|--|---|-----------------------------|--|---------|-------------------------------------|-------------------------------|--------|
| Abay | Rainfed production | 1 | 2 | 400 | 80 | 10 | 0.1 | 40 | b5734636-5db9-489a-8f-fursa | Fursa |
| Qaqaba | Rainfed production | 1 | 1 | 200 | 40 | 2 | 0.02 | 4 | 5b78b836-aea8-47f4-b1-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 0.25 | 50 | 1000 | 25 | 0.25 | 12.5 | da8a4283-618f-4d4f-81-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 200 | 65 | 10 | 0.1 | 20 | cc17475b-8961-4184-b5-fursa | Fursa |
| Abay | Rainfed production | 1 | 0.5 | 100 | 30 | 20 | 0.2 | 20 | cc17475b-8961-4184-b5-fursa | Fursa |
| Danka | Rainfed production | 1 | 1 | 200 | 4000 | 25 | 0.25 | 50 | 1caa907e-9c0d-43a2-8c-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 1.5 | 400 | 5500 | 20 | 0.2 | 80 | 40ac808b-4e0d-494c-8f-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 1.5 | 400 | 58 | 15 | 0.15 | 60 | f7dd3723-12e4-43ae-91-fursa | Fursa |
| Kan biroo (ibsi) | Rainfed production | 1 | 1.5 | 400 | 5800 | 50 | 0.5 | 200 | f7dd3723-12e4-43ae-91-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 300 | 40 | 3 | 0.03 | 9 | 9b421476-9666-4a34-bf-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 3 | 800 | 8000 | 50 | 0.5 | 300 | 563e91d7-fd22-4ee2-9a-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 250 | 4000 | | 0 | 0 | 396e91d7-c130-47d2-ae-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 2 | 600 | 6500 | 45 | 0.45 | 270 | 0f9fa217-0ffc-4009-832f-fursa | Fursa |
| Danka | Rainfed production | 1 | 1 | 200 | 3000 | 16 | 0.16 | 32 | e3808d1a-72d8-417d-8f-fursa | Fursa |
| Kan biroo (ibsi) | Rainfed production | 1 | 0.5 | 100 | 2500 | 7 | 0.07 | 7 | b399d18d-d175-4730-a1-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 0.5 | 55 | 1400 | 75 | 0.75 | 41.25 | 09c1f6a0-ed01-4596-ae-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 1 | 240 | 38 | 26.315 | 0.26315 | 63.156 | 9e3c573a-1995-480e-9d-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 0.5 | 100 | 20 | | 0 | 0 | 330712e3-618f-4be9-8d-fursa | Fursa |
| Qubsa | Rainfed production | 1 | 0.5 | 100 | 1600 | 25 | 0.25 | 25 | a83a63ca-2cb9-4d0c-8a-fursa | Fursa |

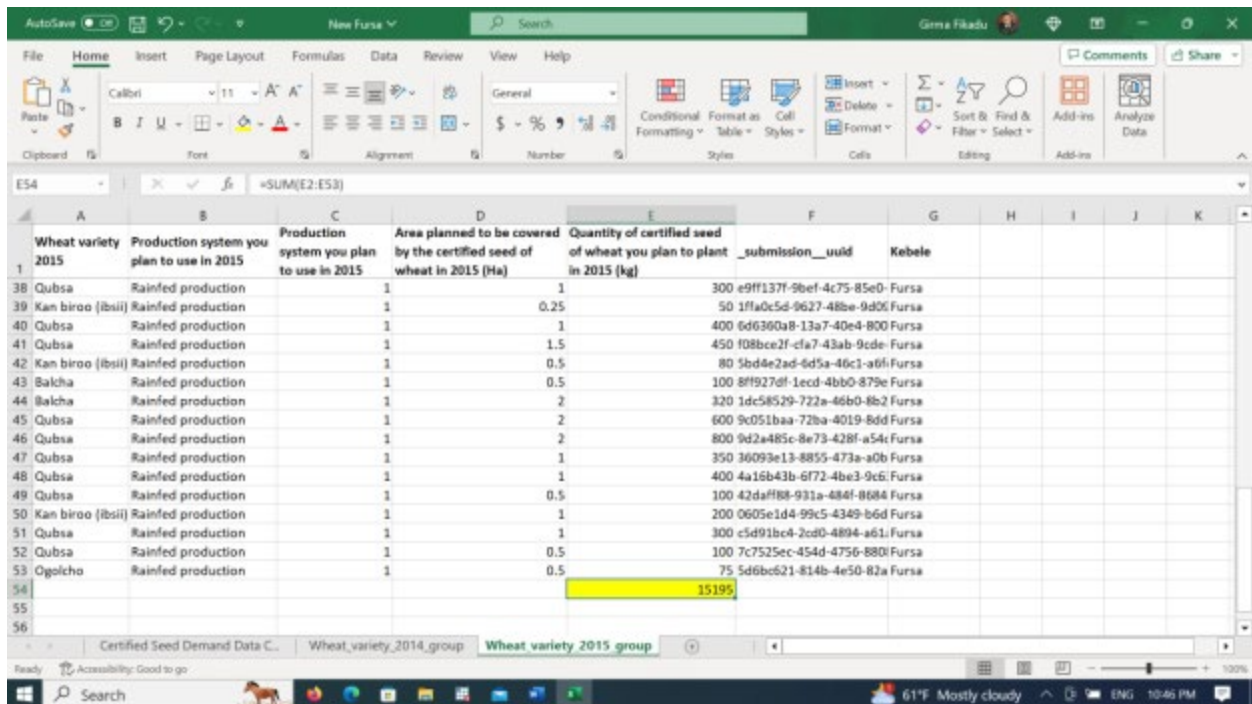
By calculating the total amount covered from own saving and calculating the percentage share of own saved seed from the grand total gives the following figures.



| Wheat variety in 2014 | Production system used in 2014 | Production system used in 2014 | Area covered by the certified seed of wheat in 2014 | Quantity of certified seed of wheat planted | Quantity of wheat harvested | Percent of wheat seed was used from own saving | Decimal | Amount of seed used from own saving | _submission__uuid | Kebele |
|-----------------------|--------------------------------|--------------------------------|---|---|-----------------------------|--|---------|-------------------------------------|------------------------------|--------|
| 36 Qubsa | Rainfed production | 1 | 1 | 350 | 4500 | 35 | 0.35 | 122.5 | 6d6360a8-13a7-40e4-8f Fursa | |
| 37 Qubsa | Rainfed production | 1 | 2 | 650 | 8000 | 38 | 0.38 | 247 | f08bce2f-cfa7-43ab-9cd Fursa | |
| 38 Danika | Rainfed production | 1 | 2 | 650 | 8500 | 41 | 0.41 | 266.5 | f08bce2f-cfa7-43ab-9cd Fursa | |
| 39 Shorima | Irrigated production | 0 | | | | 100 | 1 | 0 | 5bd4e2ad-6d5a-46c1-a1 Fursa | |
| 40 Danika | Rainfed production | 1 | 0.5 | 100 | 1500 | 33 | 0.33 | 33 | 8ff927df-1ecd-4bb0-87f Fursa | |
| 41 Danika | Rainfed production | 1 | 0.75 | 150 | 3000 | | 0 | 0 | 08ac3c2a-c555-4f74-8bf Fursa | |
| 42 Qubsa | Rainfed production | 1 | 0.75 | 150 | 4000 | 13 | 0.13 | 19.5 | 08ac3c2a-c555-4f74-8bf Fursa | |
| 43 Danika | Rainfed production | 1 | 1 | 160 | 3000 | | 0 | 0 | 1dc58529-722a-46b0-8f Fursa | |
| 44 Qubsa | Rainfed production | 1 | 1 | 400 | 4500 | 38 | 0.38 | 152 | c02f5953-91c2-47f8-907 Fursa | |
| 45 Qubsa | Rainfed production | 1 | 2.5 | 950 | 9000 | 45 | 0.45 | 427.5 | 9c051baa-72ba-4019-8c Fursa | |
| 46 Danika | Rainfed production | 1 | 2.5 | 950 | 8000 | 42 | 0.42 | 399 | 9c051baa-72ba-4019-8c Fursa | |
| 47 Qubsa | Rainfed production | 1 | 2 | 500 | 8500 | 37 | 0.37 | 185 | 9d2a485c-8e73-428f-a5 Fursa | |
| 48 Qubsa | Rainfed production | 1 | 1 | 350 | 4500 | 45 | 0.45 | 157.5 | 36093e13-8855-473a-a Fursa | |
| 49 | | | | 12920 | | | | 3724.546 | | |
| 50 | | | | | | Percentage from own saved seed | 28.8 | | | |

Meaning, considering the total amount of wheat seed demand in the kebele is 12,920 kg of which 3,724.5 kg is covered from own saving.

Therefore, 28.8 % of the total seed demand for Fursa kebele is covered from own saving. Since we now have the percentage covered from own saving, we now proceed to calculate seed demand for 2015 EC, based on the seed demand data collected for 2015 EC.



| Wheat variety 2015 | Production system you plan to use in 2015 | Production system you plan to use in 2015 | Area planned to be covered by the certified seed of wheat in 2015 (Ha) | Quantity of certified seed of wheat you plan to plant in 2015 (kg) | _submission__uuid | Kebele |
|----------------------|---|---|--|--|--------------------------------|--------|
| 38 Qubsa | Rainfed production | 1 | 1 | 300 | e9ff137f-9bef-4c75-85e0- Fursa | |
| 39 Kan biroo (lbsii) | Rainfed production | 1 | 0.25 | 50 | 1fa0c5d-9627-48be-9d06 Fursa | |
| 40 Qubsa | Rainfed production | 1 | 1 | 400 | 6d6360a8-13a7-40e4-800 Fursa | |
| 41 Qubsa | Rainfed production | 1 | 1.5 | 450 | f08bce2f-cfa7-43ab-9cde Fursa | |
| 42 Kan biroo (lbsii) | Rainfed production | 1 | 0.5 | 80 | 5bd4e2ad-6d5a-46c1-a6f1 Fursa | |
| 43 Balcha | Rainfed production | 1 | 0.5 | 100 | 8ff927df-1ecd-4bb0-879e Fursa | |
| 44 Balcha | Rainfed production | 1 | 2 | 320 | 1dc58529-722a-46b0-8b2 Fursa | |
| 45 Qubsa | Rainfed production | 1 | 2 | 600 | 9c051baa-72ba-4019-8dd Fursa | |
| 46 Qubsa | Rainfed production | 1 | 2 | 800 | 9d2a485c-8e73-428f-a54 Fursa | |
| 47 Qubsa | Rainfed production | 1 | 1 | 350 | 36093e13-8855-473a-a0b Fursa | |
| 48 Qubsa | Rainfed production | 1 | 1 | 400 | 4a16b43b-6f72-4be3-9c6 Fursa | |
| 49 Qubsa | Rainfed production | 1 | 0.5 | 100 | 42daff88-931a-484f-8684 Fursa | |
| 50 Kan biroo (lbsii) | Rainfed production | 1 | 1 | 200 | 0605e1d4-99c5-4349-b6d Fursa | |
| 51 Qubsa | Rainfed production | 1 | 1 | 300 | c5d91bc4-2cd0-4894-a61 Fursa | |
| 52 Qubsa | Rainfed production | 1 | 0.5 | 100 | 7c7525ec-454d-4756-880 Fursa | |
| 53 Ogoicho | Rainfed production | 1 | 0.5 | 75 | 5d6bc621-814b-4e50-82a Fursa | |
| 54 | | | | 15195 | | |

The above data indicates that 15,195 kg of wheat seed is demanded in Fursa kebele.

This figure is the total demand for 52 sampled households in the kebele.

- On average, 292.2 kg (=15195 kg/52) of wheat seed is demanded by each wheat producer in the kebele.
- Based on the 2014 data indicated above, 28.8% of the wheat seed demand amount is covered from own saving.
- By subtracting this amount, which is 84.2 (28.8%), at individual level, 208 kg is demanded from different sources.
- From this figure, we can calculate the total seed demand for Fursa from different sources. We do this by multiplying the net seed demand at individual level from different sources (excluding the amount covered from own saving) by the total number of wheat producing farmers in the kebele, which is 283 farmers.
- Accordingly, the total among of seed demanded from different sources for Fursa kebele in 2015 is 58,864 (=208*283)

In effect, the seed demand should be aggregated for each crop and variety. The aggregation should be done for each kebele. The aggregation done at kebele level determine the aggregation at woreda and then zone level.

Group Exercise #3: Form 3 groups to conduct SUM of area and seed demand at kebele, woreda and zone levels using seed data collected in Arsi

REPORTING

The seed demand data aggregated at kebele level is reported to woreda. Each woreda aggregates the seed demand data reported by all kebeles under it and report to the zone. Likewise, the zone aggregates the data reported to woredas under it and report to the regional bureau of agriculture. Below are illustrations of the summary of seed demand data taking example from Lode Hetosa woreda

i) Kebele level data:

- Seed demand database is maintained **per household**; for each crop and for each variety
- The demand data usually covers area planned to plant with improved seed and quantity of seed demanded.
- Augmentation data: seed demand covers area and quantity of seed along with saved seed which is used to adjust the quantity of seed newly demanded
- Moreover, trend data is established.
- The summary data to be reported to woreda is SUM of HH level area and quantity of seed by crop and variety:

Woreda: Lode Hetosa; Kebele: Fursa, Male= ---- Female=-----Total=-----

| Crop | Variety | 2014EC | | 2015EC | | 2016EC | |
|------|---------|----------|---|----------|---|----------|--|
| | | Area(ha) | Quantity of Certified Seed Planted (kg) | Area(ha) | Quantity of Certified Seed Planted (kg) | Area(ha) | Quantity of Certified Seed Demanded (kg) |
| | | | | | | | |

| | | | | | | | |
|-------|--------------|-----------|---------------|-----------|--------------|-----------|--------------|
| Wheat | Qubsa | 10 | 6,840 | 6 | 1,100 | | |
| Wheat | Danka | 4 | 780 | 5 | 1,110 | 1 | 250 |
| Wheat | Abay | | | 5 | 900 | 5 | 1,010 |
| Wheat | Shorima | 1 | 208 | 1 | 150 | 13 | 2,620 |
| Wheat | Boru | 3 | 490 | 20 | 3,880 | | |
| Wheat | Ogolcho | 4 | 3,215 | | | | |
| Wheat | Shaki | | | 3 | 550 | 4 | 750 |
| Wheat | Shina | | | 1 | 100 | | |
| Wheat | Total | 22 | 11,533 | 39 | 7,790 | 23 | 4,630 |

ii) Woreda level aggregation:

- Seed demand database is maintained per kebele; for each crop and for each variety
- The demand data usually covers area planned to plant with improved seed and quantity of seed demanded.
- Moreover, trend data is established.
- The summary data to be reported to Zone is SUM of kebele level area and quantity of seed by crop and variety:

Woreda: Lode Hetosa; No. of Kebeles: _____ No. of HHs: Male: _____ Female: _____

| Crop | Variety | 2014EC | | 2015EC | | 2016EC | |
|-------|------------------|-------------|---|--------------|---|-------------|--|
| | | Area(ha) | Quantity of Certified Seed Planted (kg) | Area(ha) | Quantity of Certified Seed Planted (kg) | Area(ha) | Quantity of Certified Seed Demanded (kg) |
| Wheat | Qubsa | 45.1 | 13,985.0 | 35.5 | 7,000.0 | | |
| Wheat | Danka | 8.1 | 1,660.0 | 12.5 | 2,010.0 | 5.3 | 1,050.0 |
| Wheat | Ogolcho | 22.1 | 6,900.0 | 14.3 | 2,615.0 | 7.3 | 1,330.0 |
| Wheat | Balcha | | | 11.5 | 2,155.0 | 19.3 | 3,660.0 |
| Wheat | Abay | 2.5 | 500.0 | 10.8 | 1,750.0 | 14.8 | 2,960.0 |
| Wheat | Shorima | 1.0 | 208.0 | 2.0 | 400.0 | | |
| Wheat | Liben | 1.3 | 230.0 | 1.8 | 365.0 | 1.3 | 265.0 |
| Wheat | Boru | 5.3 | 940.0 | 27.8 | 5,605.0 | 15.3 | 3,020.0 |
| Wheat | Qubsa | | | | | 24.5 | 5,180.0 |
| Wheat | Shaki | | | 2.8 | 550.0 | 3.5 | 750.0 |
| Wheat | Shina | 0.3 | 40.0 | 0.5 | 100.0 | | |
| Wheat | Sub Total | 85.6 | 24,463.0 | 119.3 | 22,550.0 | 91.1 | 18,215.0 |

iii) Zone level seed demand data aggregation

- Seed demand database is maintained per woreda; for each crop and for each variety
- The demand data usually covers area planned to plant with improved seed and quantity of seed demanded.
- Moreover, trend data is established.

- The summary data to be reported to Region is SUM of Woreda level area and quantity of seed by crop and variety:

Zone: Arsi; No. of Woreda: _____ No. of HHs: Male: _____ Female: _____

| Crop | Variety | 2014 EC | | 2015 EC | | 2016 EC | |
|--------------|------------------|--------------|---|--------------|---|--------------|--|
| | | Area(ha) | Quantity of Certified Seed Planted (kg) | Area(ha) | Quantity of Certified Seed Planted (kg) | Area(ha) | Quantity of Certified Seed Demanded (kg) |
| Wheat | Qubsa | 137.4 | 42,310.0 | 150.5 | 38,800.0 | 104.0 | 26,542.0 |
| Wheat | Qaqaba | 1.0 | 200.0 | | | | |
| Wheat | Danka | 97.9 | 21,064.5 | 24.0 | 6,030.0 | 14.5 | 3,110.0 |
| Wheat | Ogolcho | 22.6 | 7,100.0 | 18.3 | 3,415.0 | 10.3 | 2,010.0 |
| Wheat | Wane | 2.0 | 400.0 | 1.0 | 200.0 | 42.8 | 8,918.5 |
| Wheat | Balcha | 8.8 | 1,977.5 | 36.0 | 7,196.0 | | |
| Wheat | Abay | 5.8 | 1,210.0 | 60.8 | 11,975.0 | 84.8 | 20,267.5 |
| Wheat | Shorima | 2.3 | 508.0 | 3.0 | 600.0 | 5.0 | 1,420.0 |
| Wheat | Liben | 16.3 | 4,683.0 | 20.5 | 4,190.0 | 14.0 | 2,865.0 |
| Wheat | Hidase | | | | | 3.0 | 900.0 |
| Wheat | Digalu | 1.0 | 200.0 | | | | |
| Wheat | Boru | 8.3 | 1,700.0 | 40.8 | 8,195.0 | 24.5 | 4,820.0 |
| Wheat | Ga'ambo | | | 0.3 | 50.0 | | |
| Wheat | Dambal | 1.5 | 300.0 | | | | |
| Wheat | Danda'a | 0.5 | 100.0 | | | | |
| Wheat | Dursa | 1.8 | 450.0 | | | | |
| Wheat | Shaki | 1.3 | 300.0 | 23.5 | 5,070.0 | 27.0 | 5,837.5 |
| Wheat | Shina | 0.3 | 40.0 | 0.5 | 100.0 | | |
| Wheat | Sub Total | 308.4 | 82,543.0 | 379.0 | 85,821.0 | 329.8 | 76,690.5 |

Trend analysis

Trend of seed demand



Trend of kebele level certified seed demand (kg)



Trend of woreda level certified seed demand (kg)



Some reasons for declining seed demand

- Fluctuation in area cultivated
- Increasing seed price
- Increasing fertilizer cost
- Analysis of **trend, reasons for changes** are important for **decisions for increased inputs used and hence increased productivity**

Annex 3: Training Evaluation Form

Feed the Future Global Supporting Seed Systems for Development (S34D) activity - Capacity Building Training on Data Management

We value your feedback. Please take a minute to fill out this evaluation question about the training. Please indicate your rating of the training in the categories below by circling the appropriate rating scale. In this feedback form, there are no WRONG or RIGHT answers. You do not need to put your name on this form – your responses are anonymous. Please respond to ALL the questions below to help us to improve similar future training.

Please indicate your impressions of the items listed below.

| | Strongly Agree | Agree | Neutral | Disagree | Strongly Disagree |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. The training met my expectations. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. I will be able to apply the knowledge learned. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. The training objectives for each topic were identified and followed. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. The content was organized and easy to follow. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. The materials distributed were pertinent and useful. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. The trainer was knowledgeable. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. Class participation and interaction were encouraged. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Adequate time was provided for questions and discussion. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

9. How do you rate the training overall?

Excellent Good Average Poor Very poor

10. What limitations did you observe?

11. What should be improved in future training? _____

12. Other comments? _____

Thank you for your participation!