Basic Data Entry & Analysis

Training Manual

15 July 2014

Contents

1 Introduction 3

2 Form Design 4
2.1 Question numbering 4
2.2 Questions where the answer is a number 4
2.3 Questions where the answer is multiple choice (tick one) 4
2.4 Questions where the answer is tick all that apply 4
2.5 Using codes 4
2.6 Formatting 5

3 Data Entry 6
3.1 Creating an EpiData Form 6
3.1.1 Define data 6
3.1.2 Make a data file 11
3.1.3 Apply checks 13
3.1.4 Getting the toolbar to come back 17
3.2 Entering Data 19
3.3 Exporting from EpiData 22
3.3.1 Exporting directly to Excel 22
3.3.2 Exporting to a text file then importing into Excel 27

4 Data Cleaning 37
4.1 Setting up a filter 37
4.2 Auditing data entry 44
4.3 Checking each column 48
4.4 Checking jumps 51
4.5 Deleting rows 54

5 Data Analysis 58
5.1 Creating a Pivot Table 58
5.2 Checking the codes 61
5.3 Copying and pasting from Pivot Tables 64
5.3.1 Copying to Word 64
5.3.2 Copying to Excel 67
5.4 Questions where the answer is a number 70
5.4.1 Calculating totals 70
5.4.2 Calculating averages 72
5.4.3 Calculating percentages 75
5.5 Questions where the answer is multiple choice (tick one) 76
5.5.1 Calculating totals 76
5.5.2 Calculating percentages 78
5.6 Questions where the answer is tick all that apply 82
5.6.1 Calculating totals 82
5.6.2 Calculating percentages 84
5.7 Questions with jumps 88
5.8 Showing results by Station, TA, GVH, or Village 90

6 Mapping 93
6.1 Collecting GPS coordinates for locations 93
6.1.1 Download a GPS app 93
6.1.2 Go to the locations and record the GPS coordinates 93
6.2 Preparing the spreadsheet 94
6.3 Making a basic map 95

1 Introduction

This manual was created by the Child Health Program in the central region of Malawi. It gives step-by-step instructions for doing basic data entry and analysis using the following software:

Software	Use	Download link
EpiData	Data entry for surveys and forms. Exporting raw data sets.	http://www.epidata.dk/download.php *Choose EpiData Manager*
Microsoft Excel	Analysing data using Pivot Tables.	N/A
Tableau Public	Creating interactive maps and charts.	http://www.tableausoftware.com/public/download

This manual is for complete beginners. It should be used with on-the-job training by someone who has experience with data entry and analysis. The manual alone is not enough for someone to learn all the skills required.

This manual does not cover statistical analysis. For basic statistical tests (t-test, chi-square, etc) you could use the free package PSPP, available to download here: http://www.gnu.org/software/pspp/

2 Form Design

Good data entry and analysis starts with a good form design. Use the following tips to make sure your surveys / forms are well designed for data entry and analysis:

2.1 Question numbering

All questions must have a number, and the numbers must be in the same format (e.g. Q1, Q2, Q3 or A, B, C or i,ii,iii). Sub-questions can be used if necessary, and should always be in the same format. For example:

Q1. *A)*

B)

C)

2.2 Questions where the answer is a number

For these types of questions put a _____ where the answer should go. For example:

Q1. What is your age? _____ years

2.3 Questions where the answer is multiple choice (tick one)

For multiple choice questions all the answers should be listed. Each answer should have a code number. A ○ should be placed next to each answer for the enumerator to tick. For example:

Q1. Was the child sick in the last 2 weeks?

1 ○ Yes

2 ○ No

3 ○ Unknown

2.4 Questions where the answer is tick all that apply

For these types of questions all the answers should be listed. Each answer should have a code number. A ☐ should be placed next to each answer. The instructions should say “tick all that apply”. For example:

Q2. What symptoms did the child have? (tick all that apply)

1 ☐ Fever

2 ☐ Cough

3 ☐ Diarrhoea

2.5 Using codes

Codes should be placed at the top of the survey / report to identify where it came from. These codes should correspond to the Village, GVH, TA, Station, Enumerator, etc. Each code should be unique. Only the codes will be entered into the database. This will save time because it means you don’t have to type out the village or GVH name every time. For example:

Q1. Village name _____ *Code #*_____

Q2. GVH name _____ *Code #*_____

You should prepare a code book listing all the codes to be used. For example:

GVH codes

01 = Mdabwi

02 = Chizumba

03 = Khuzi

2.6 Formatting

Use tables to format the survey / report. Questions should appear on the left, with all answers appearing on the right. This will make data entry easier.

Separate questions using horizontal lines. Separate sections using headings with grey shading.

Example

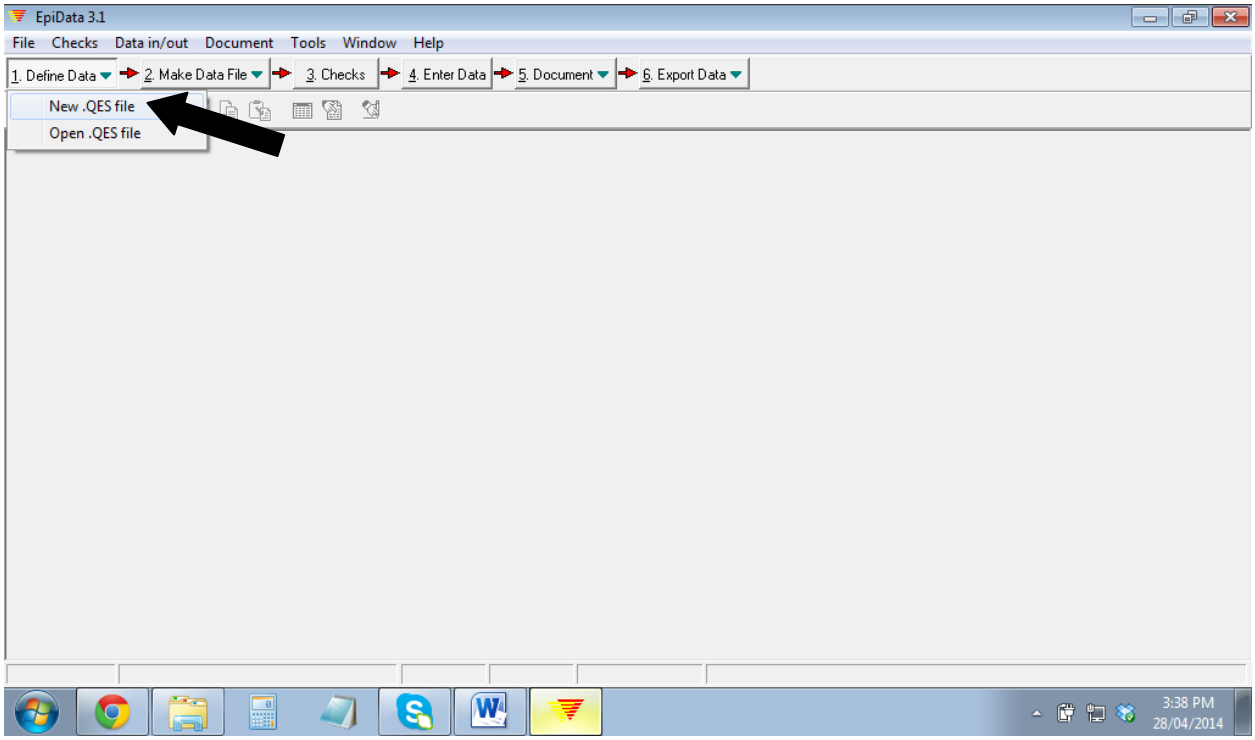
STAKEHOLDER PARTICIPATION	
1. Which members of the **pair** were present?	1 ○ Both HSA and Facilitator 2 ○ HSA only 3 ○ Facilitator only
2. How many **Village Headmen** were present?	_____
3. How many **GVHs** were present?	_____

3 Data Entry

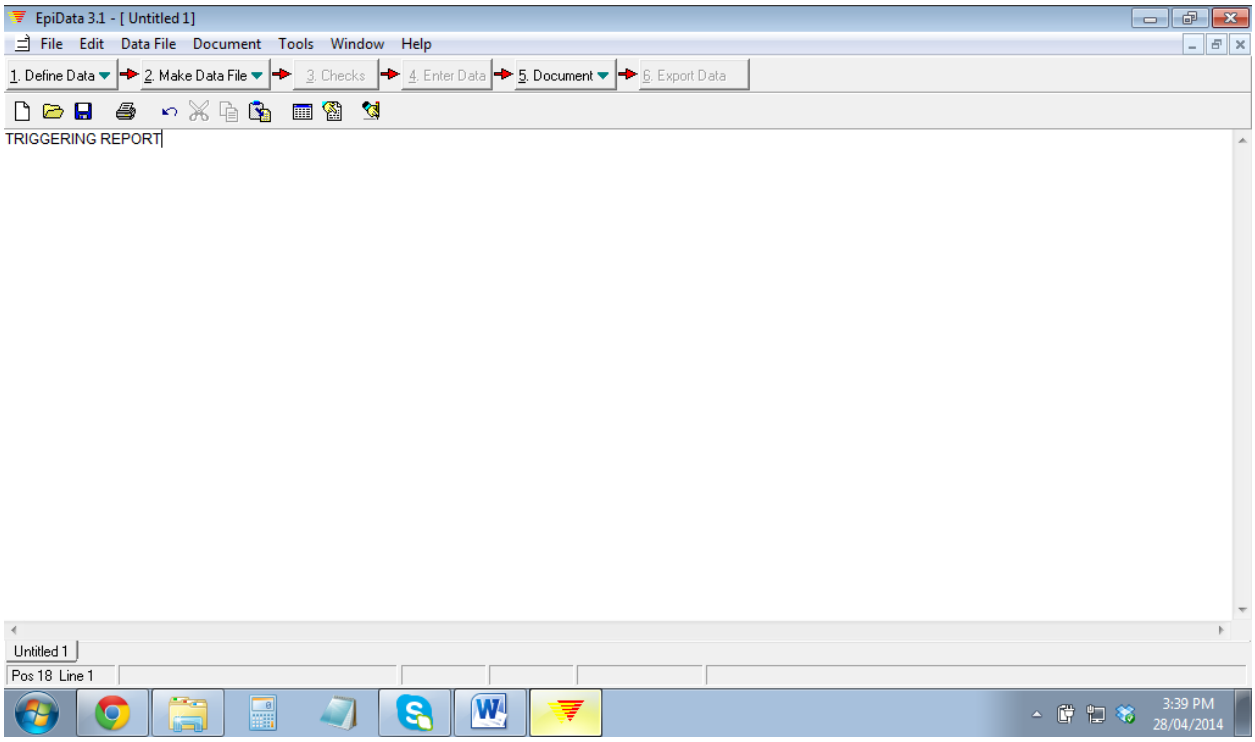
3.1 Creating an EpiData Form

3.1.1 Define data

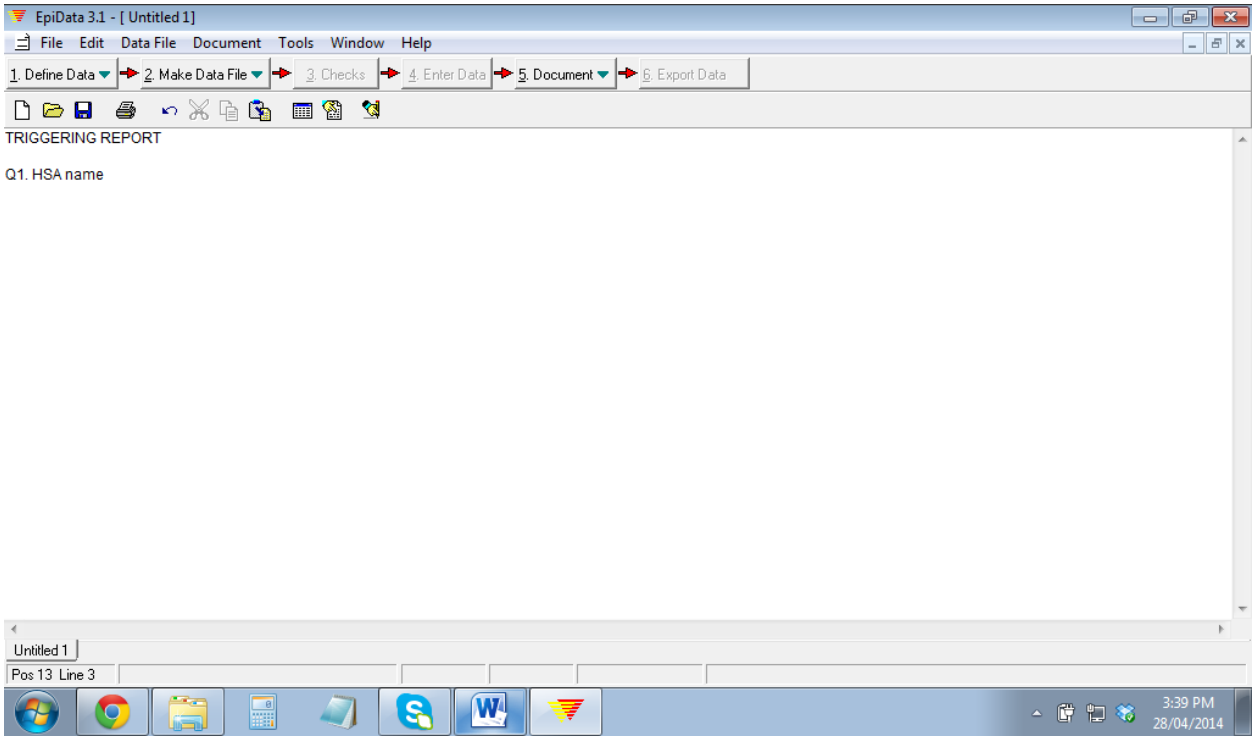
Step 1: Open EpiData. Click on the Define Data button. Choose New .QES file.



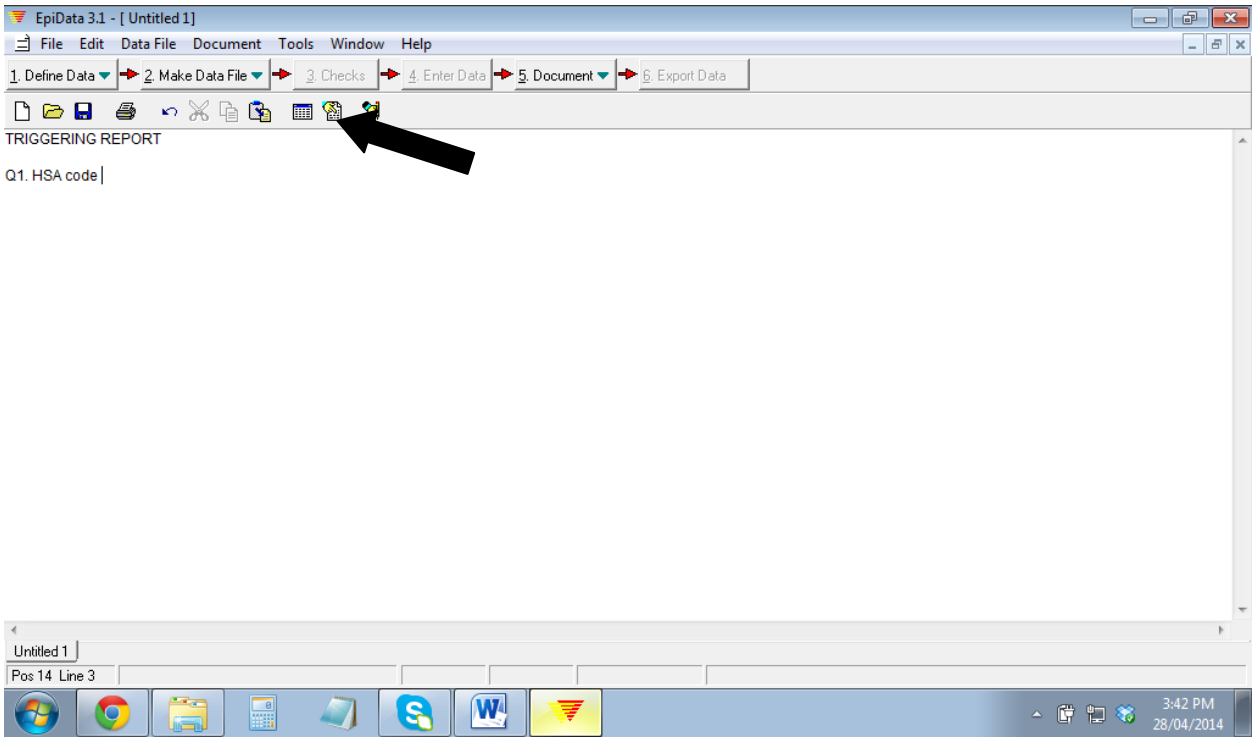
Step 2: Type a heading for the form.



Step 3: Go to a new line and type the first question.

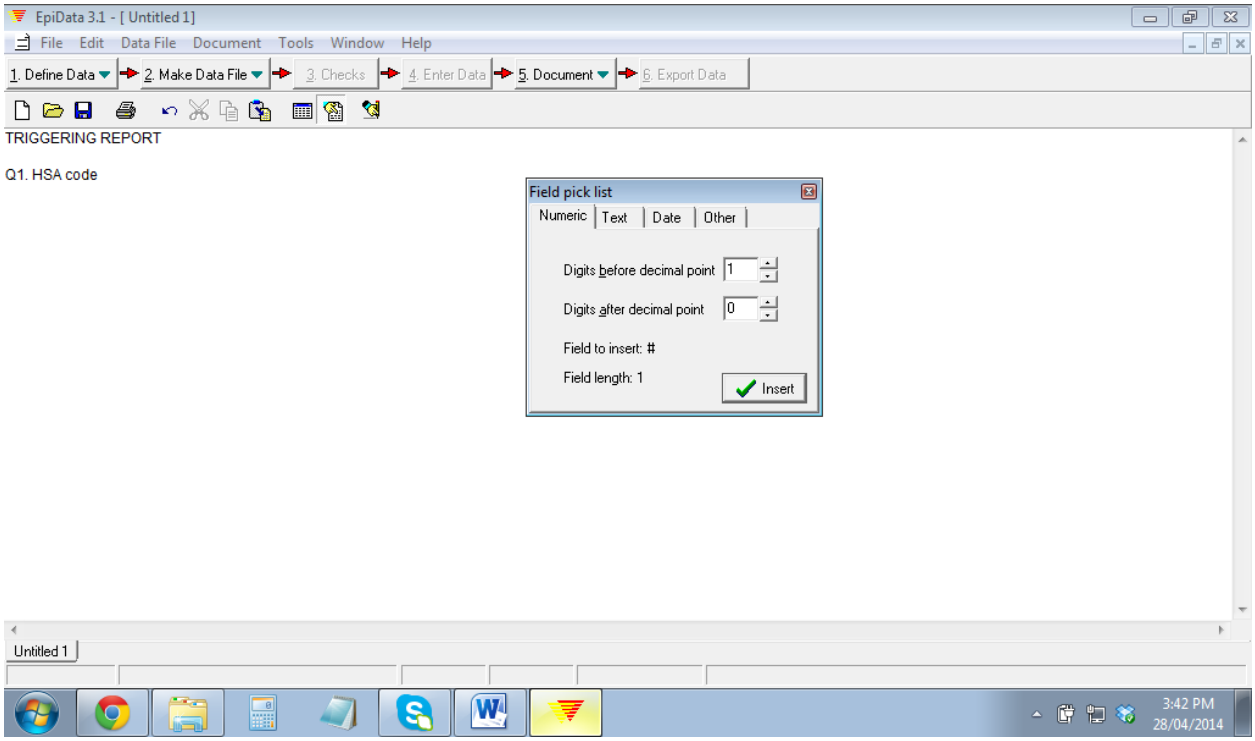


Step 4: Click on the Field Pick List icon.



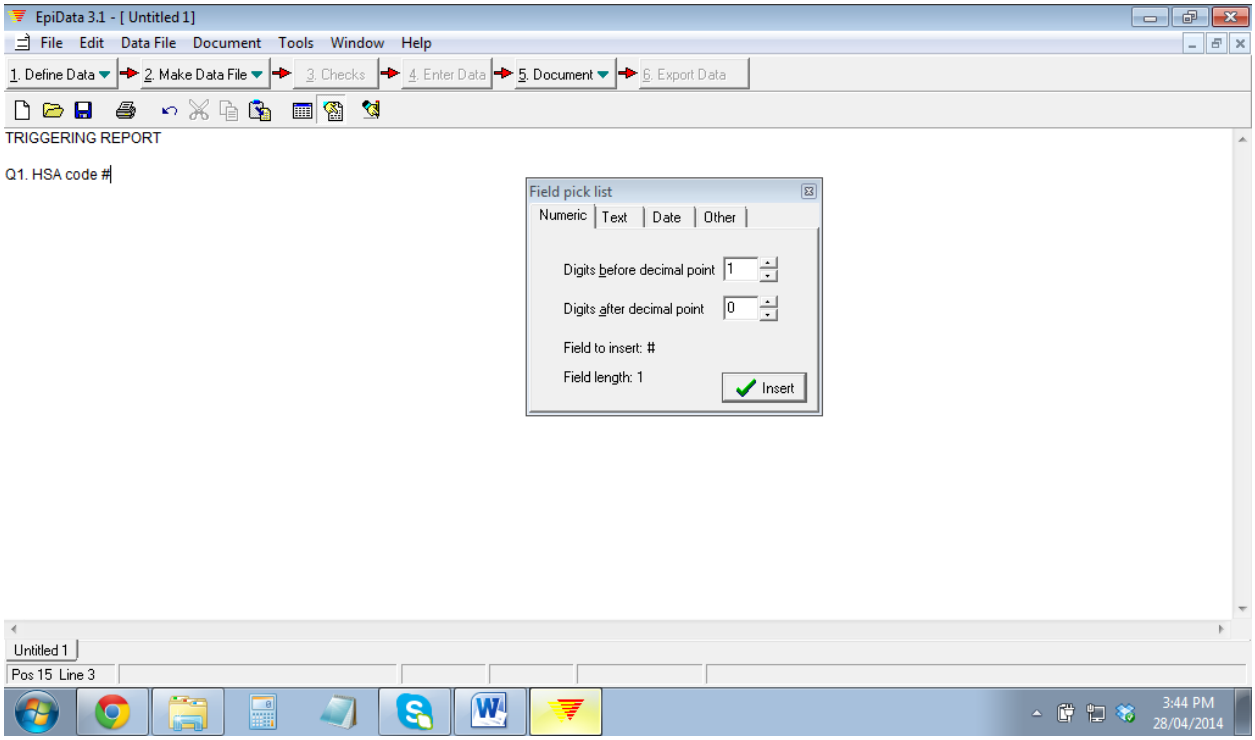
Step 5: Choose Numeric if the answer is a number, Text if it is a word, or Date if it is a date. Use the options to specify how long the field should be.

Example: For a numeric field 2014 is 4 digits, while 10 is 2 digits.



Step 6: Click Insert. A symbol will be inserted next to the question to indicate the field. The number of symbols indicates the length of the field.

Example: For a numeric field of 4 digits the symbols will appear as ####, while for 2 digits it is ##



Step 6: Follow the same steps until all question have been entered with corresponding fields.

HINTS

You can type the symbols directly to make the fields without using the pick list:

is one digit for a numeric field

_ is one letter for a text field

<dd/mm/yyyy> is a date

For multiple choice questions where you tick only one (marked by a ○) you only need to put one or two digits so that the answer number can be entered.

Example

Q14. Does the household have a latrine? 1 ○ Yes 2 ○ No

Would be entered as: Q14. Does the household have a latrine? #

For multiple choice questions where you tick more than one (marked by a ☐) you only need to put a separate question for each answer.

Example

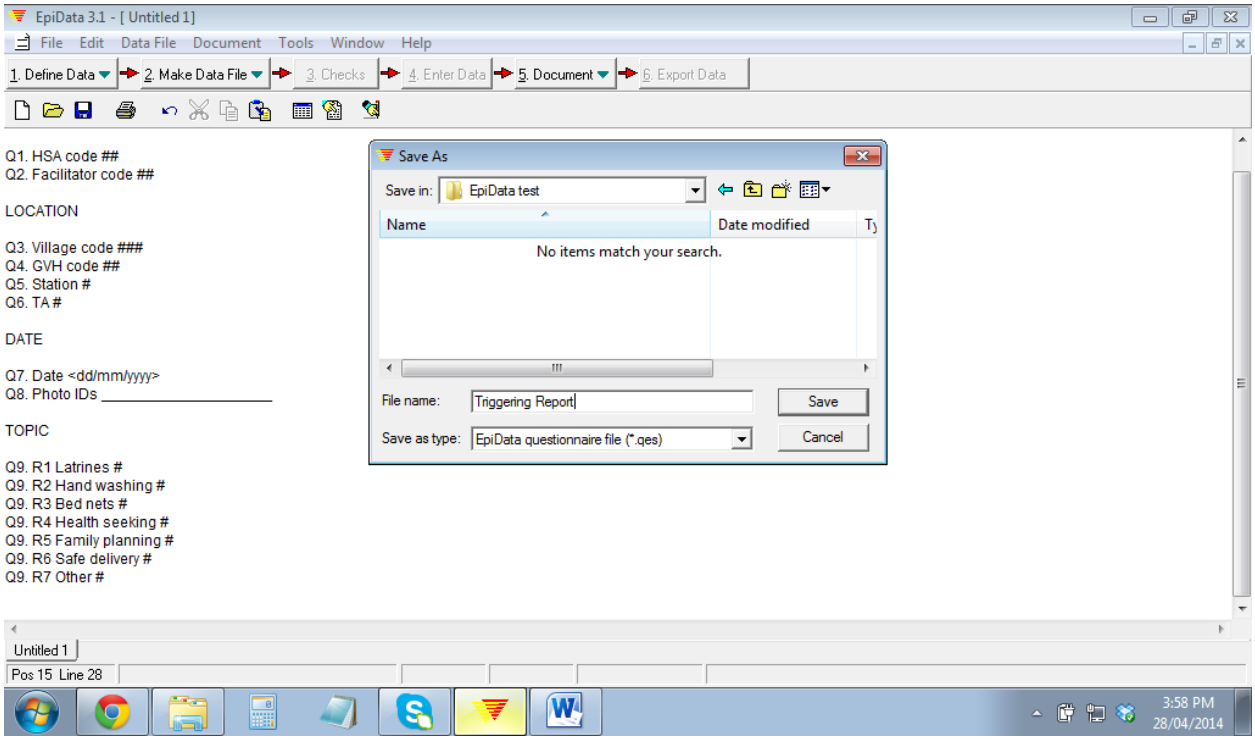
Q9. What was the topic of the activity? (tick all that were covered)

1 ☐ Latrines
2 ☐ Hand washing
3 ☐ Bed nets
4 ☐ Health seeking
5 ☐ Family planning
6 ☐ Safe delivery
7 ☐ Other

Would be entered as:

Q9.R1 Latrines #
Q9.R2 Hand washing #
Q9.R3 Bed nets #
Q9.R4 Health seeking #
Q9.R5 Family planning #
Q9.R6 Safe delivery #
Q9.R7 Other #

Step 7: Save the file.



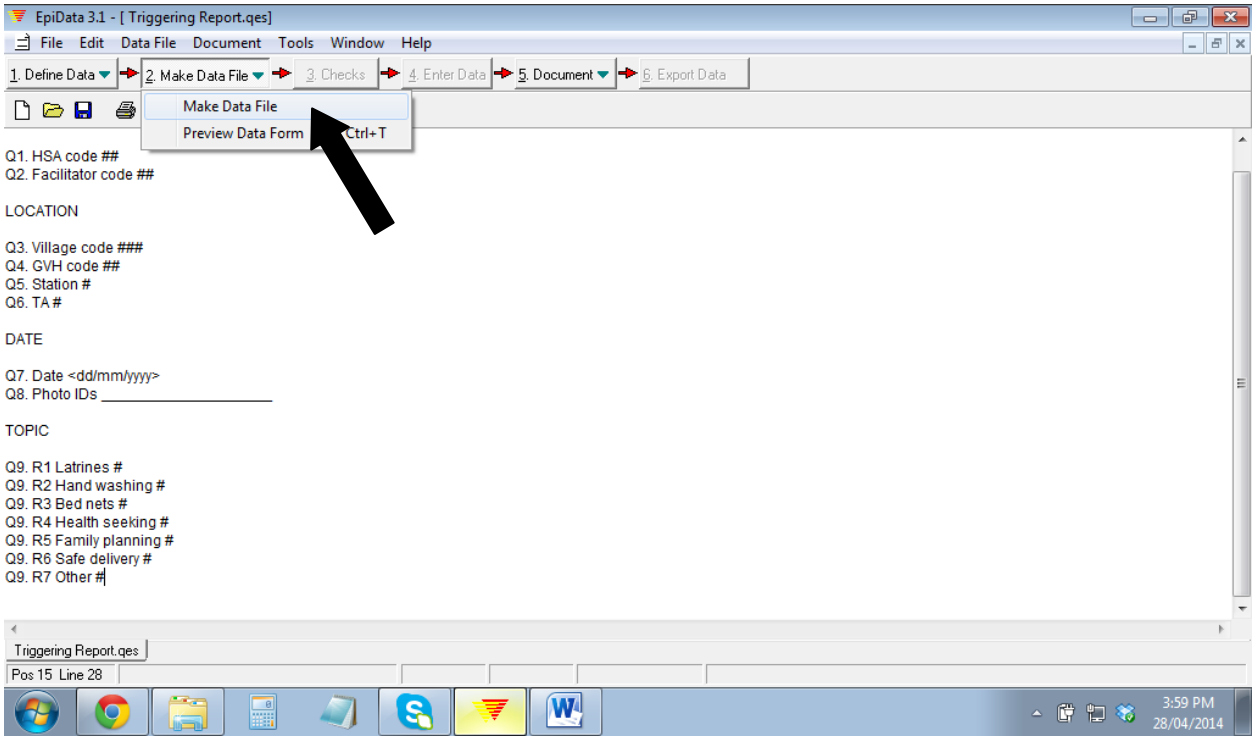
HINT

Once you start data entry you CANNOT make any changes to the questions that you typed in this step. Make sure all the questions are there before proceeding, and make sure all of them have the right number of digits. If you do come back to make changes during data entry then all the data you have entered will be DELETED.

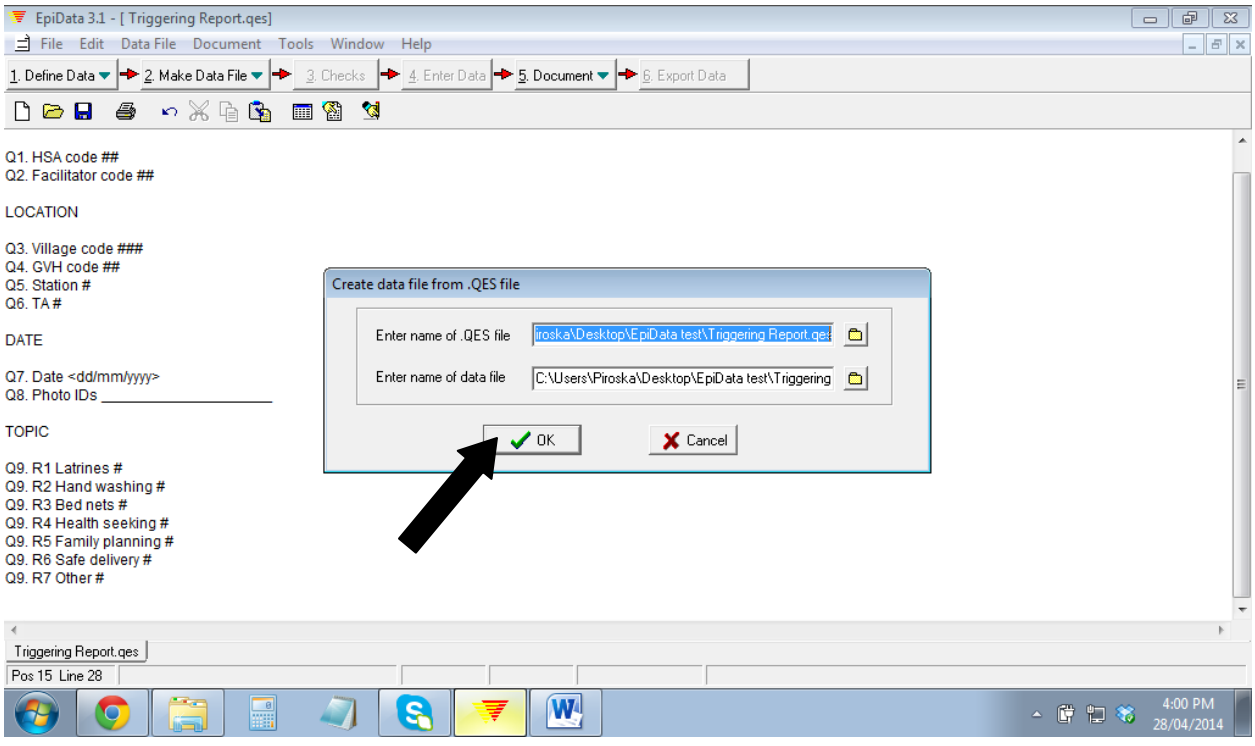
3.1.2 Make a data file

Once you have finished typing in the questions you need to make a data file.

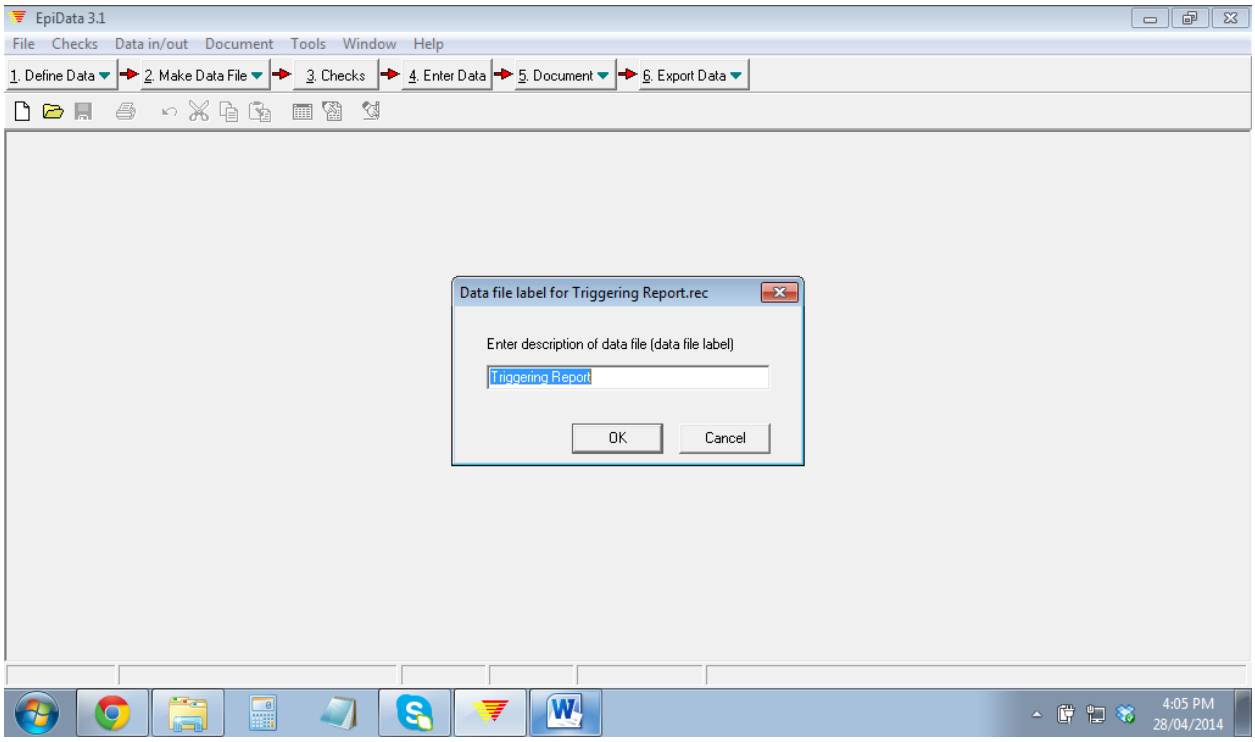
Step 1: Click on Make Data File. Choose Make Data File.



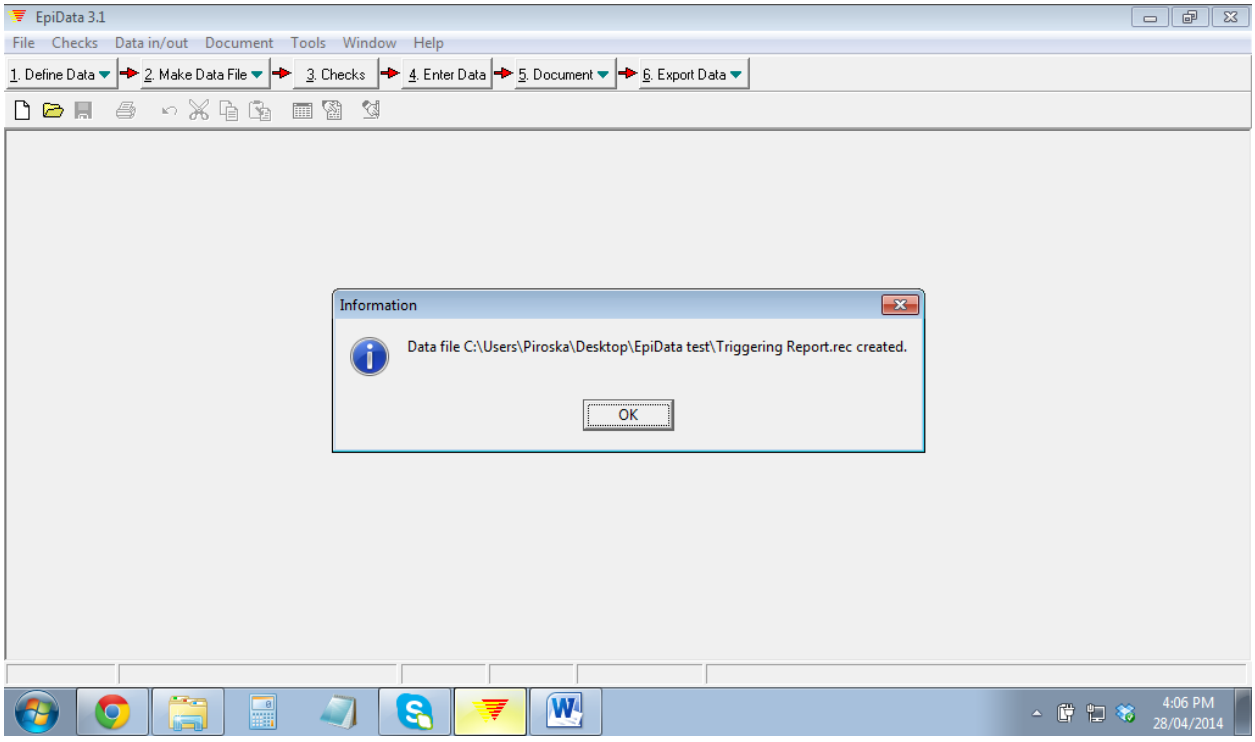
Step 2: Make sure the file name is the same as the one you were working on. Then click OK.



Step 3: Enter a name for the file. Then click OK.



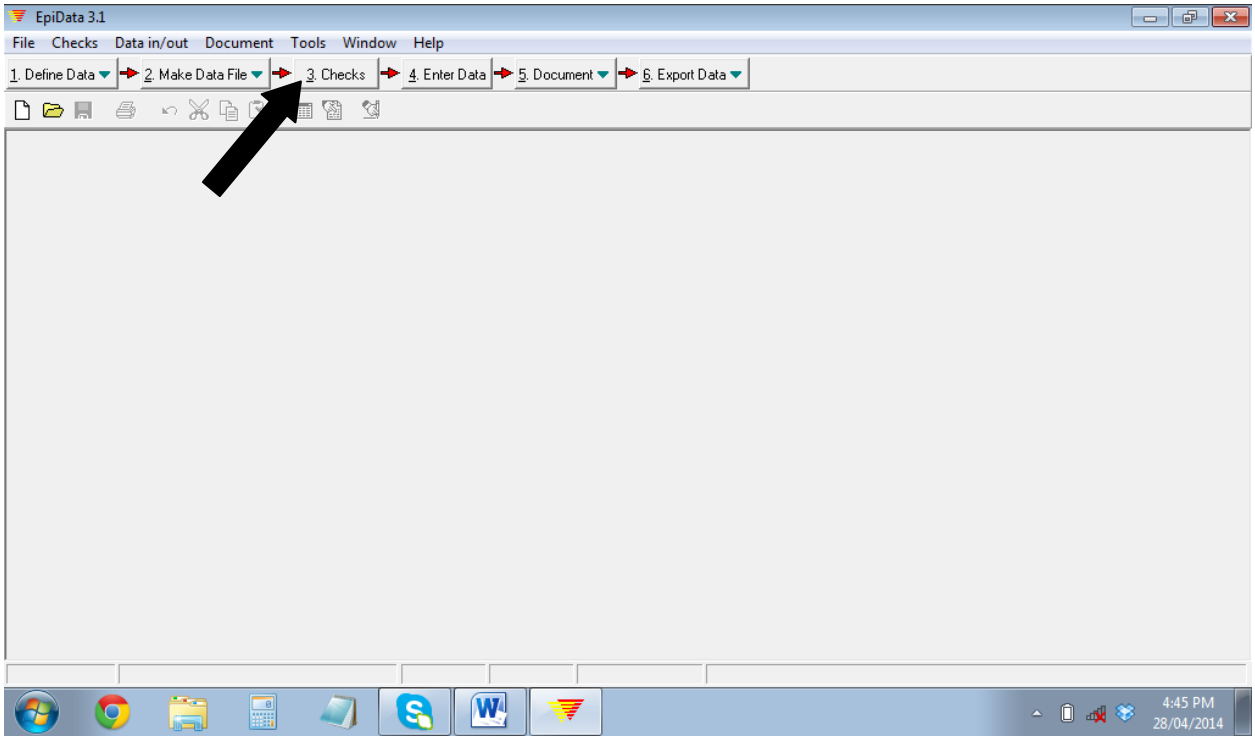
Step 4: You should receive this message to confirm it was created. Click OK.



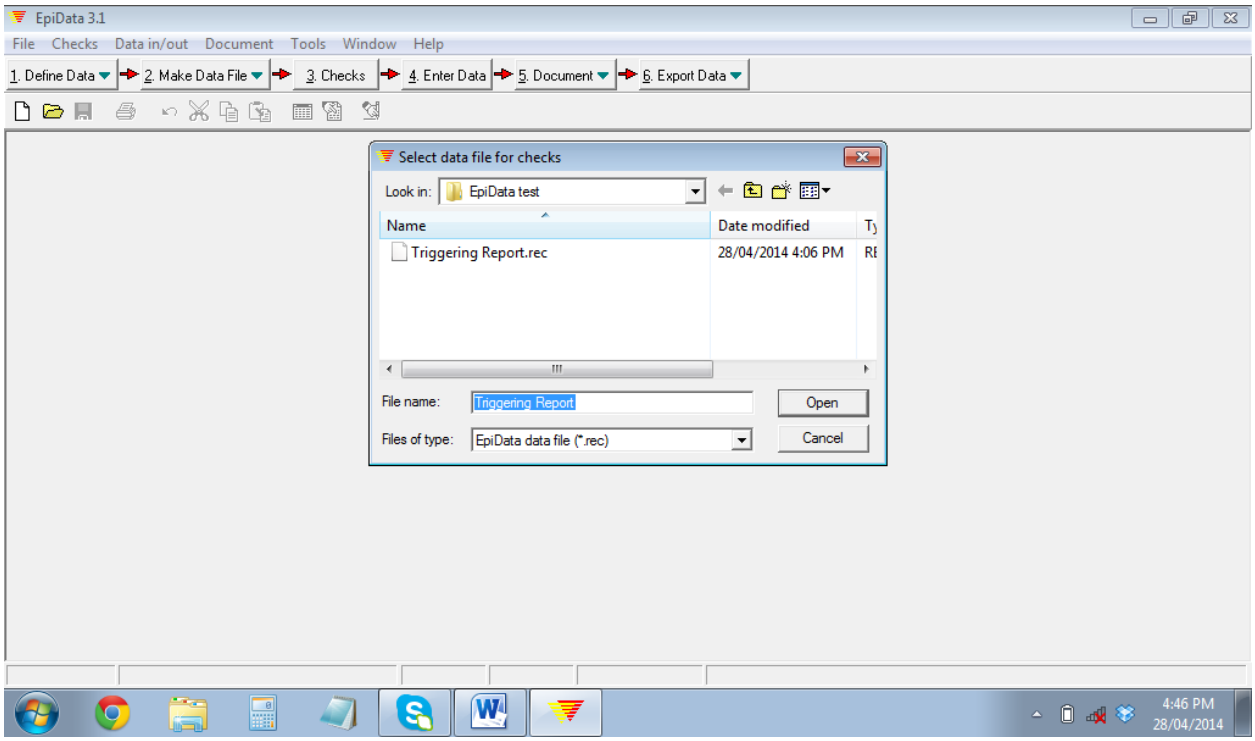
3.1.3 Apply checks

Once the data file has been created you can apply checks to restrict what is allowed to go in each field. For example, a field for Month must be between 1 and 12. You can also apply jumps so that certain fields are skipped depending on which answer you enter.

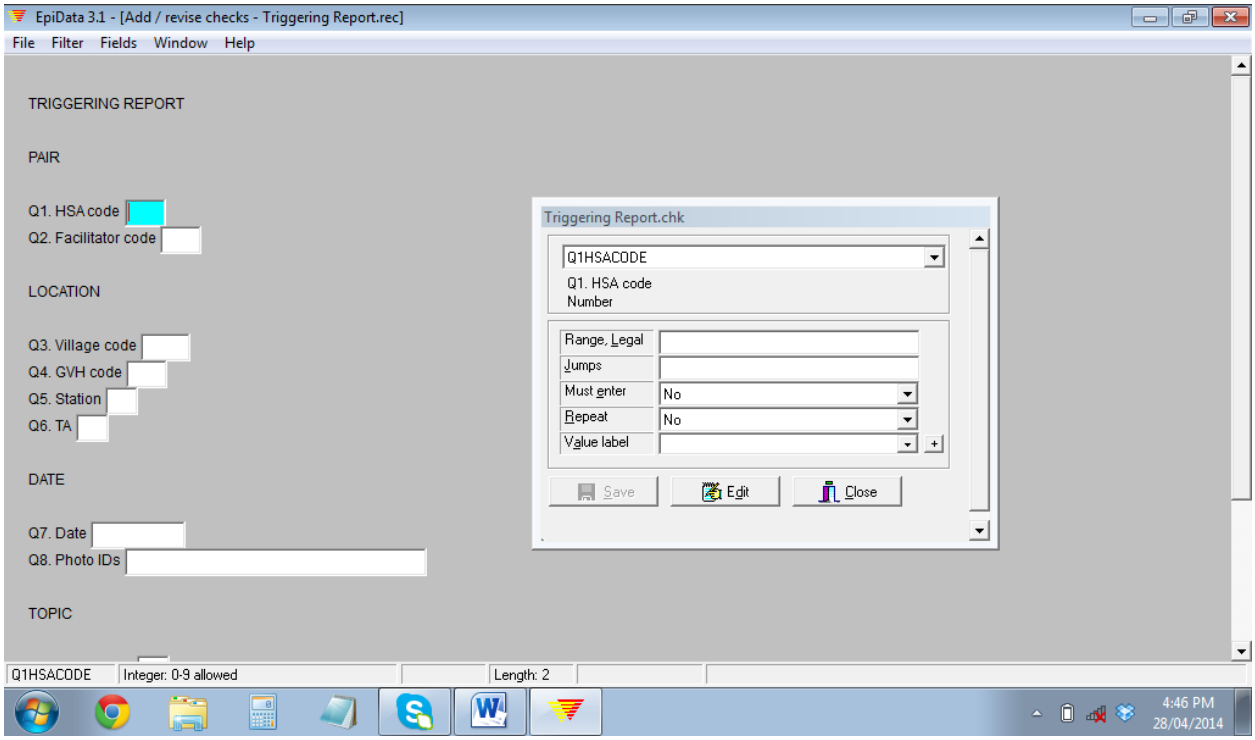
Step 1: Click on the Checks button.



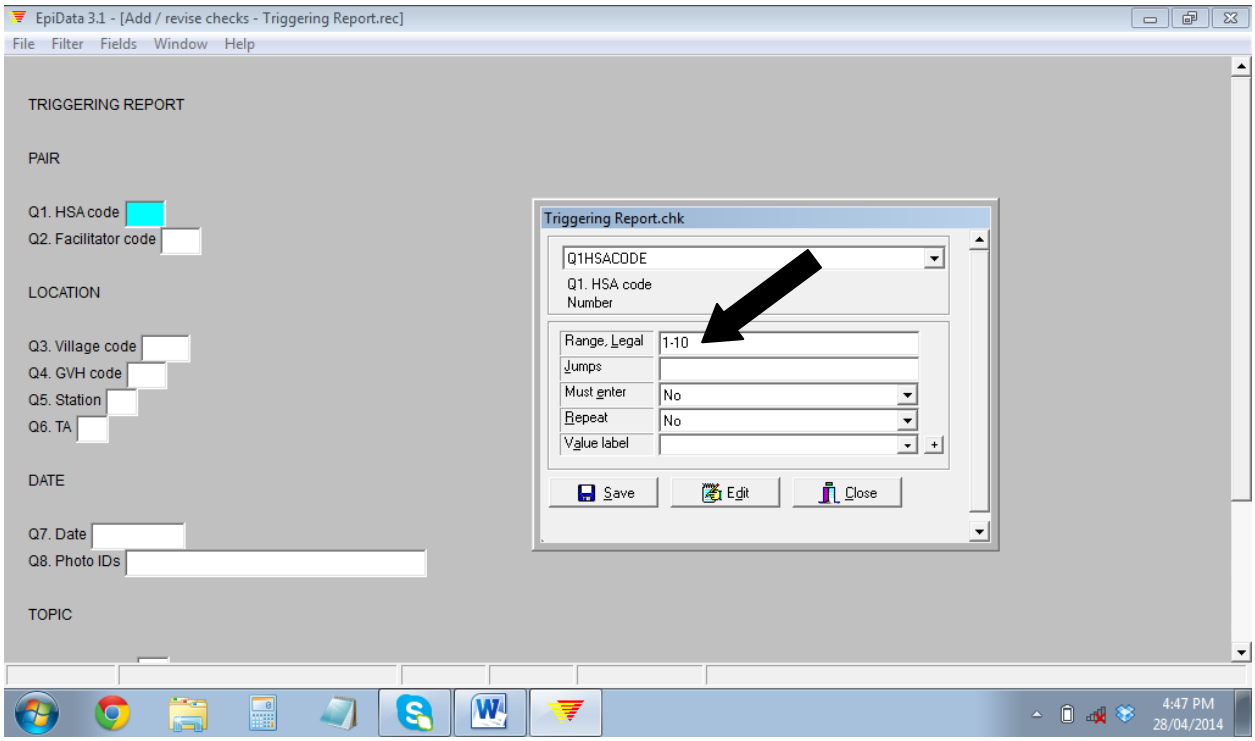
Step 2: Choose the file and click Open.



Step 3: Go to the first question.

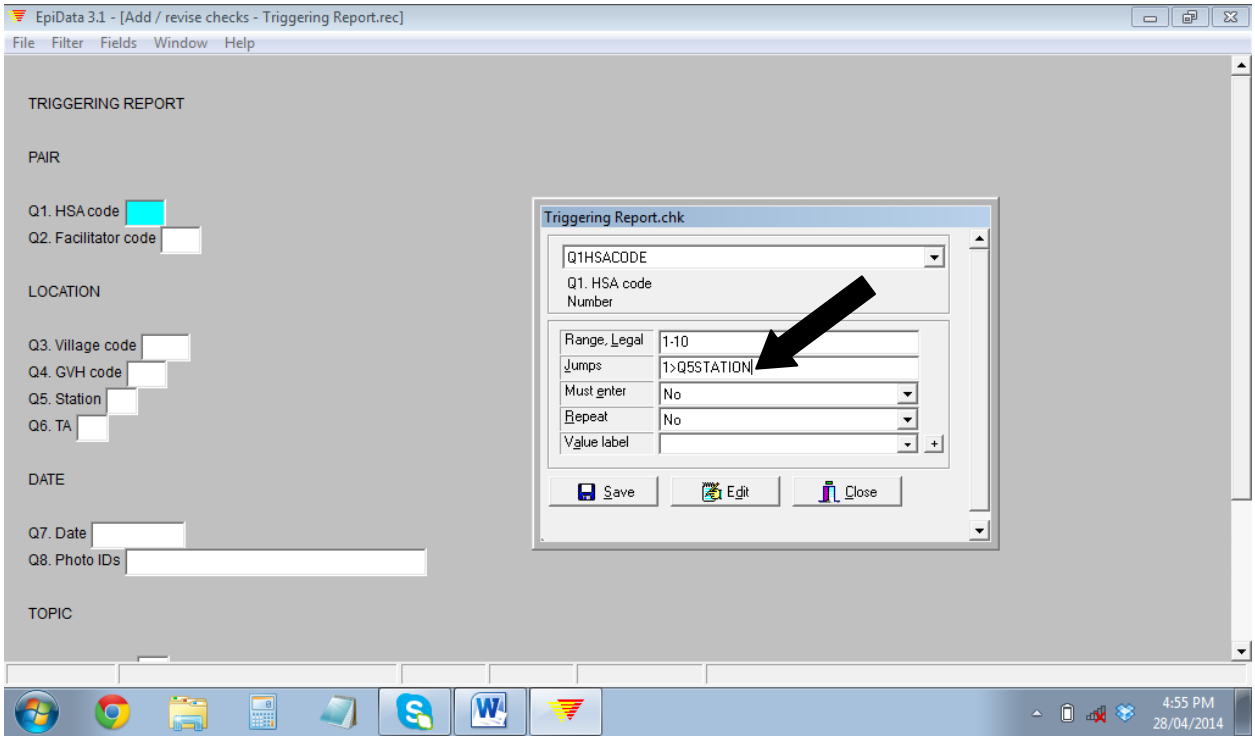


Step 4: In the Range, Legal field write the range of numbers that can be entered for that question. For example 1-10. Any number outside this range will be rejected (e.g. 20, 0)



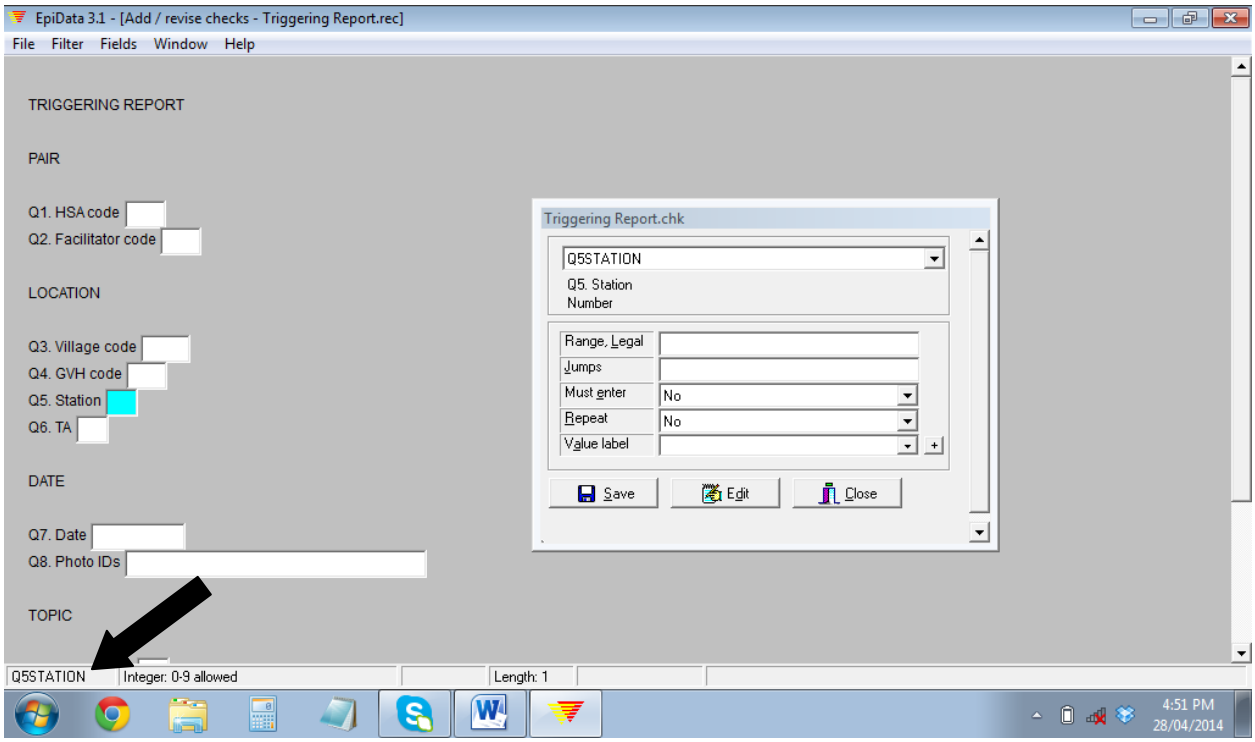
Step 5: The Jumps field can be used to skip questions depending on the answer. The format is *Answer Number > Question Name.* You must write the **full** name of the question, not just the question number.

Example: "If the answer is 1 then jump to Q5" would be written as 1> Q5STATION



HINTS:

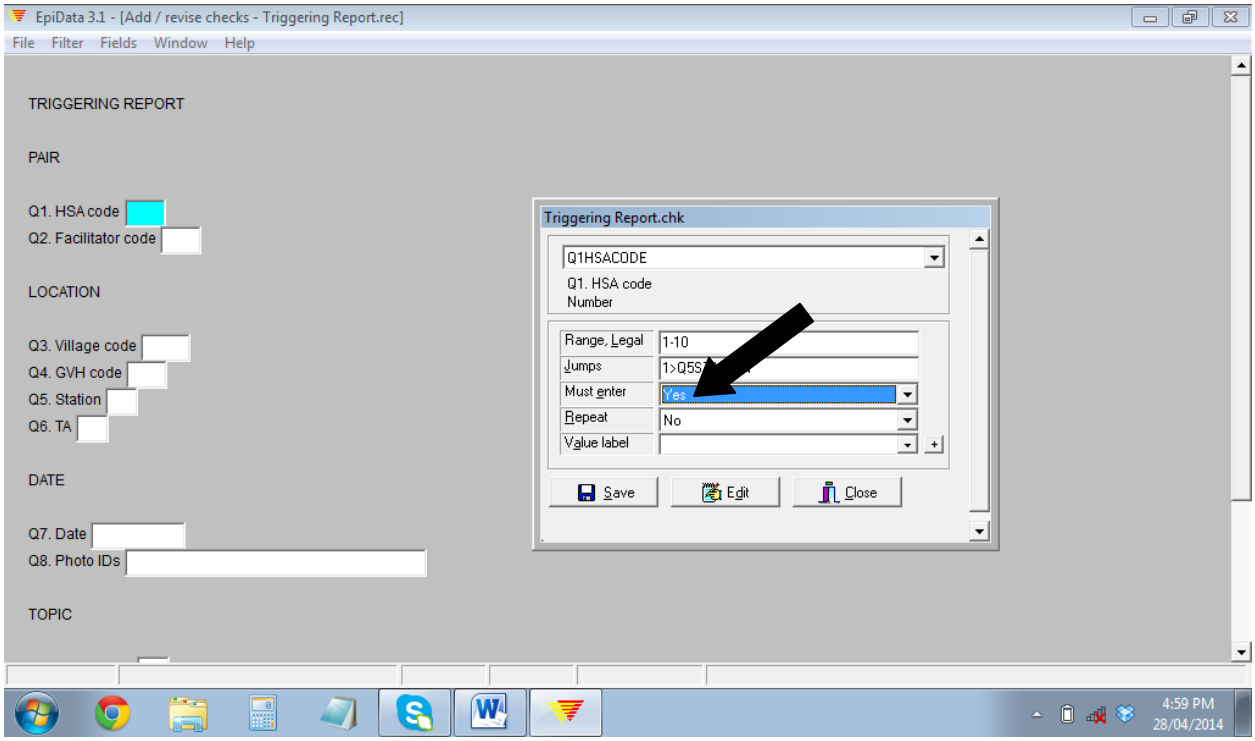
You can find the full name of a question by clicking on it and looking in the bottom left corner of the screen. In the example below the full name for Q5 is Q5STATION.



To add another jump just put a comma. For example, 1> Q5STATION, 2>Q7DATE

To jump regardless of the answer write AUTOJUMP followed by the question you want to jump to (no > is required). For example, AUTOJUMP Q5STATION

Step 6: Use the Must Enter field to force the user to complete the field. If Must Enter is set to Yes they must complete the field, if it is No then the field is optional.



Step 7: Click on the Save button. Follow the same steps for all the remaining questions in the form.

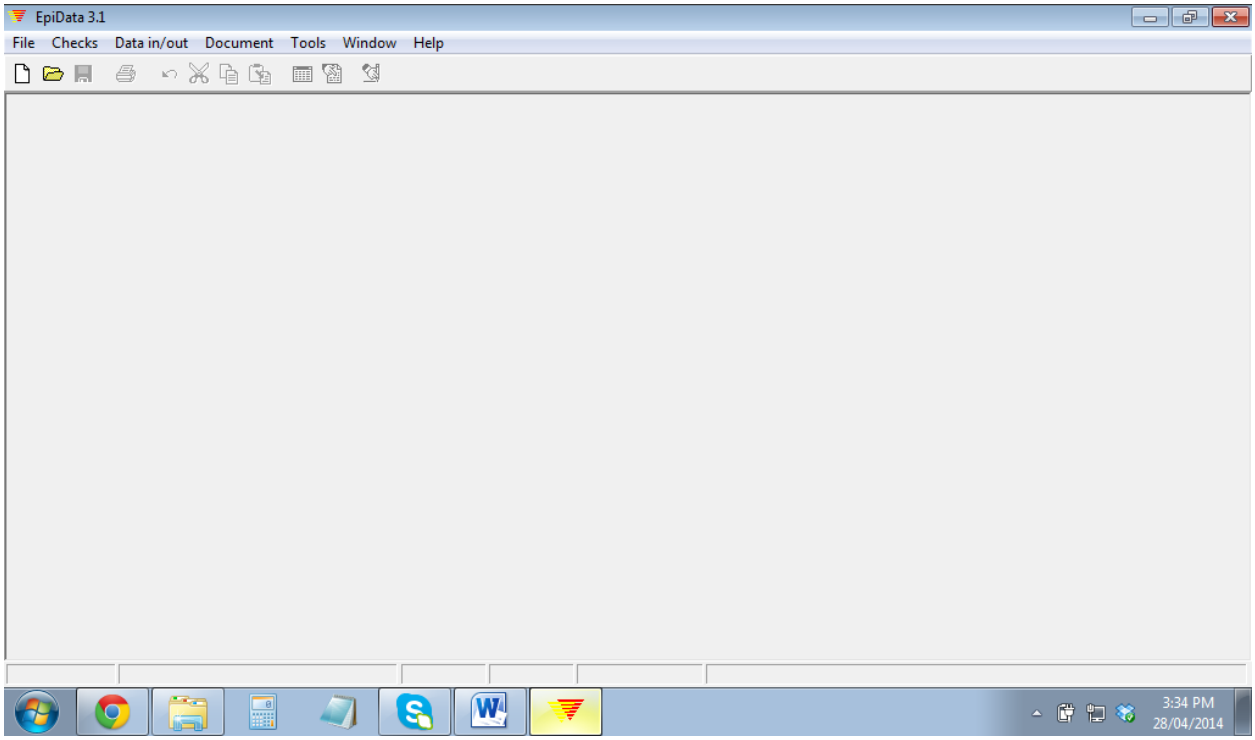
Step 8: When you are finished with all the questions click the Close button. You are now ready to enter data.

HINT

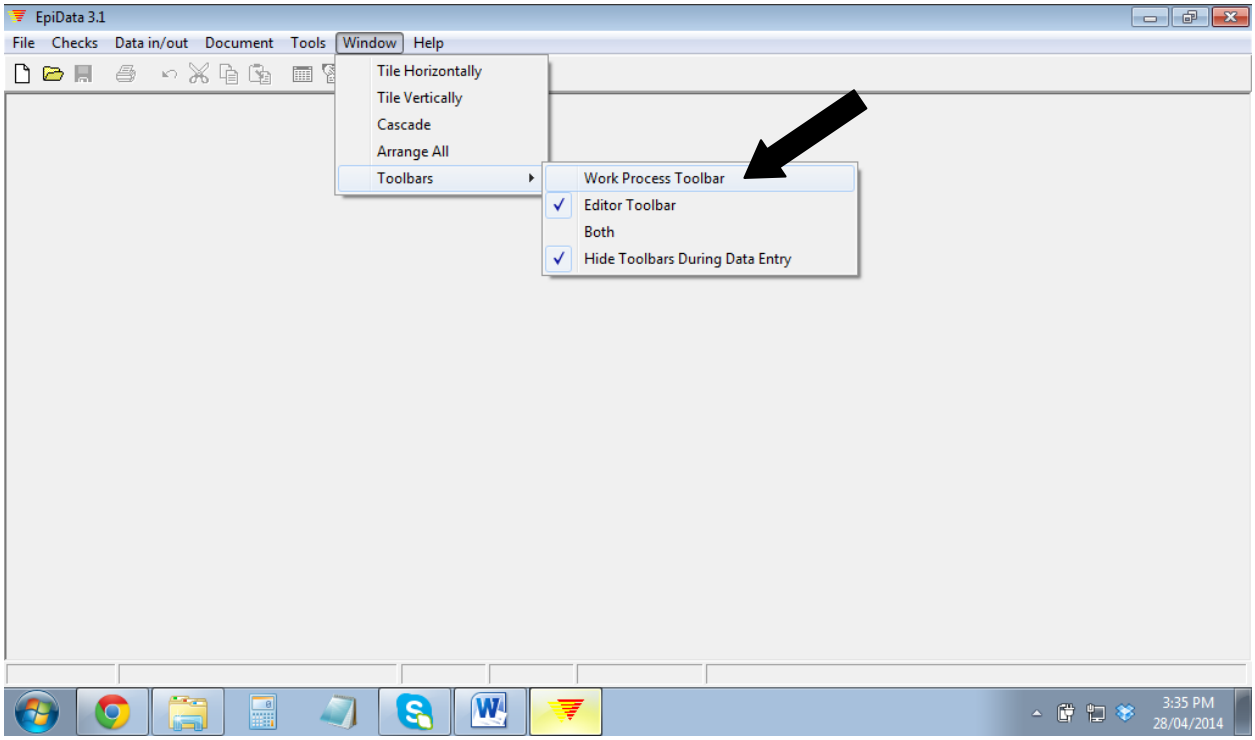
If you ever want to change a check during the data entry process just close the data entry form and come back to checks. Fix the check and click save. You can then open the data entry again and continue.

3.1.4 Getting the toolbar to come back

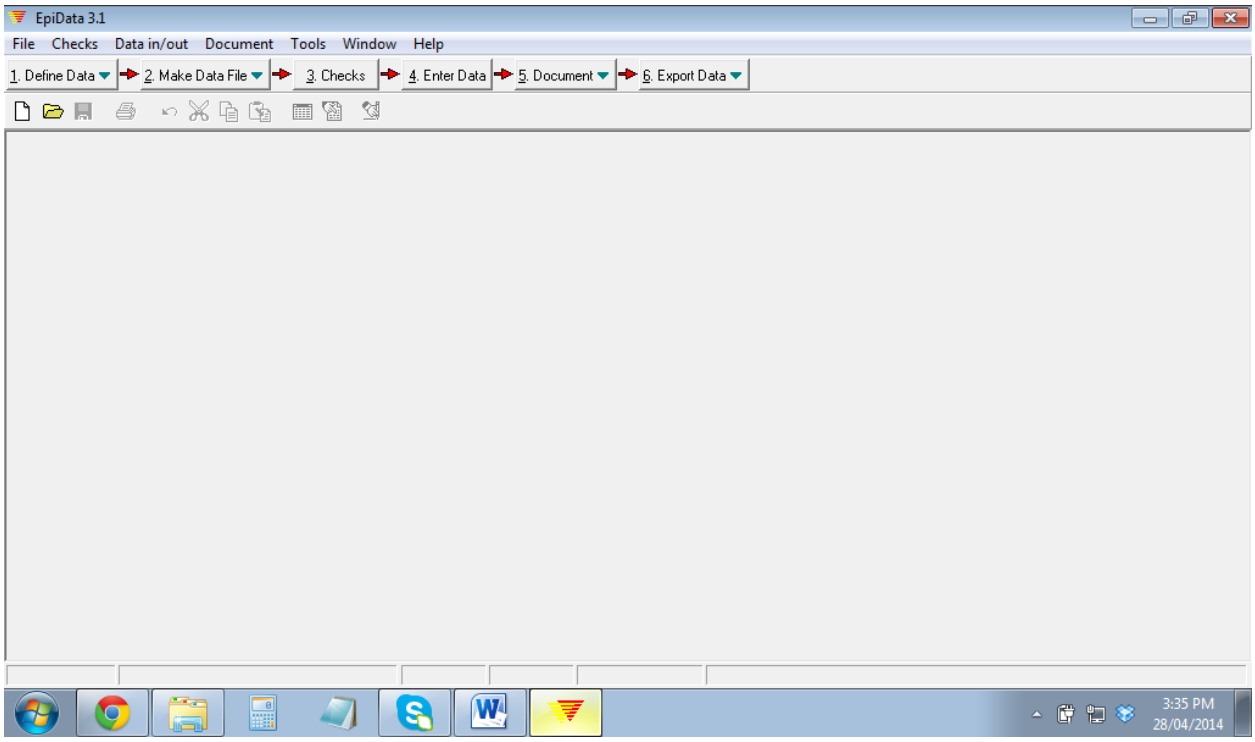
If the toolbar at the top of EpiData ever disappears (like the image below) you can get it back by following these steps.



Step 1: Go to the Window Menu. Click on Toolbars. Tick the Work Process Toolbar.



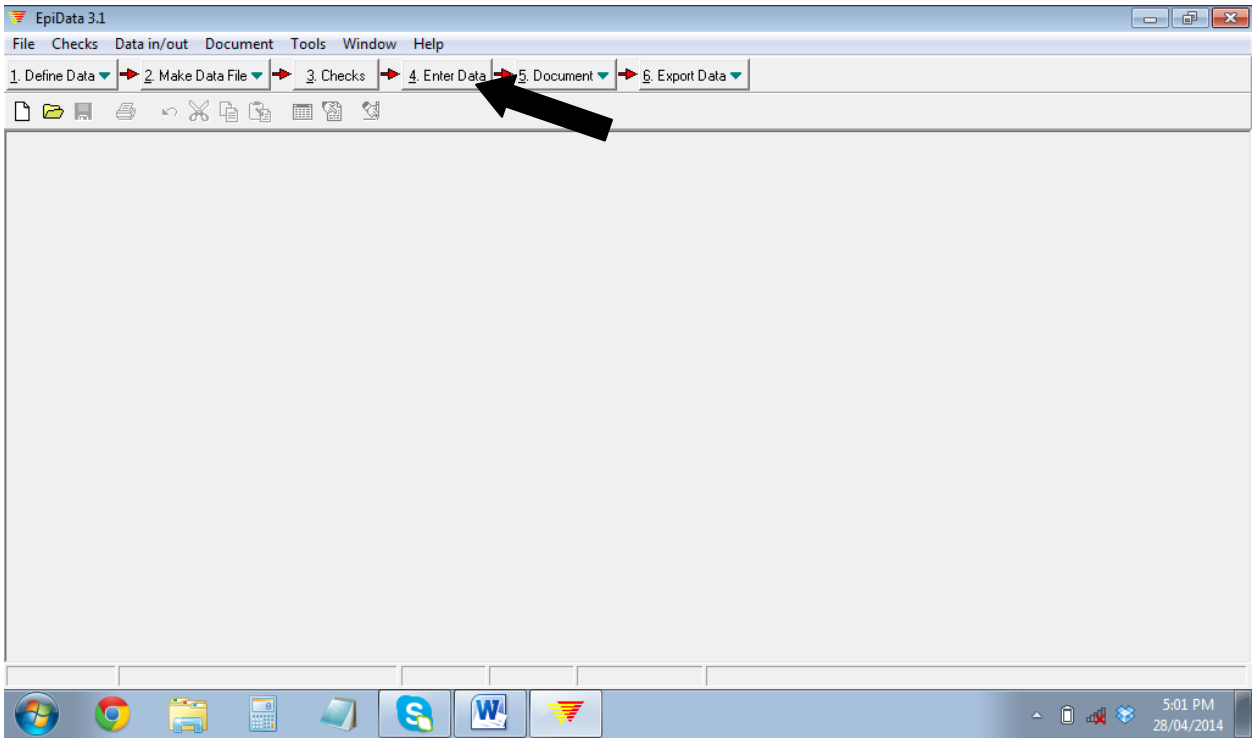
Step 2: The toolbar will appear.


EpiData 3.1
File Checks Data in/out Document Tools Window Help
1. Define Data
2. Make Data File
3. Checks
4. Enter Data
5. Document
6. Export Data
3:35 PM
28/04/2014

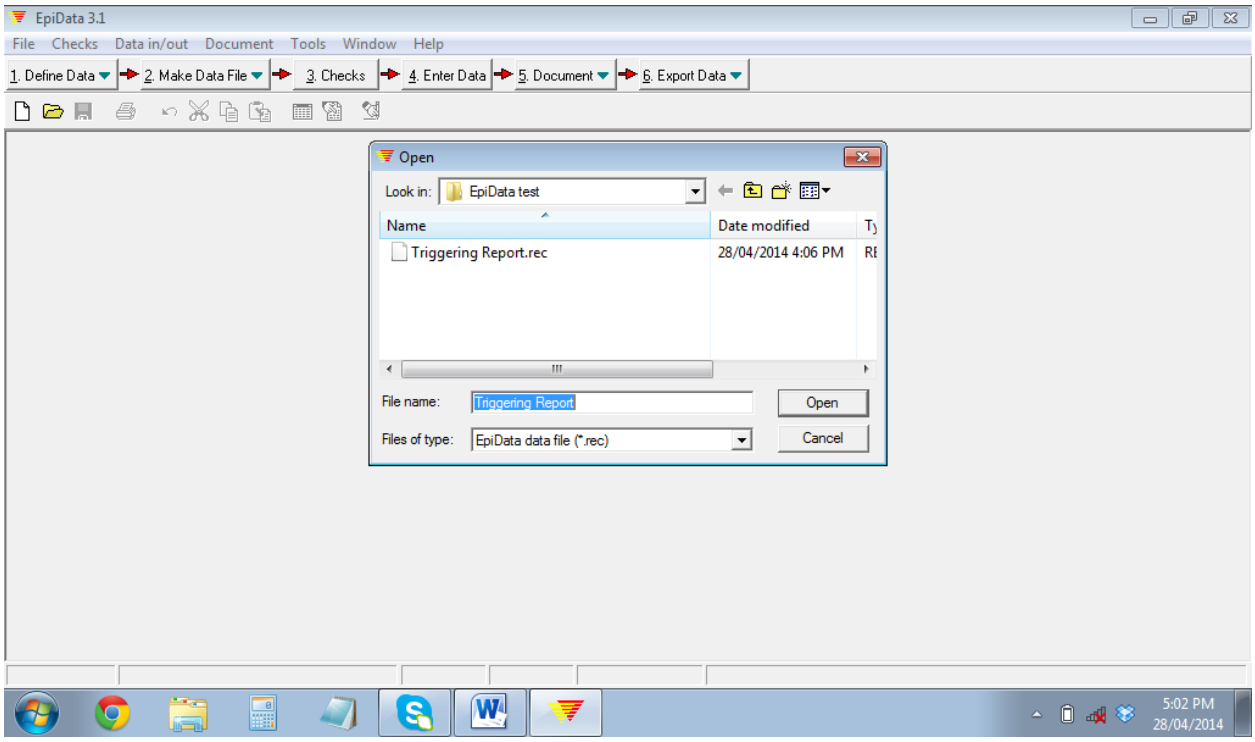
3.2 Entering Data

After you have setup the checks you are ready to enter data. First do some test data entry to make sure the form is working as you expected. If there are problems go back to the previous steps and fix them.

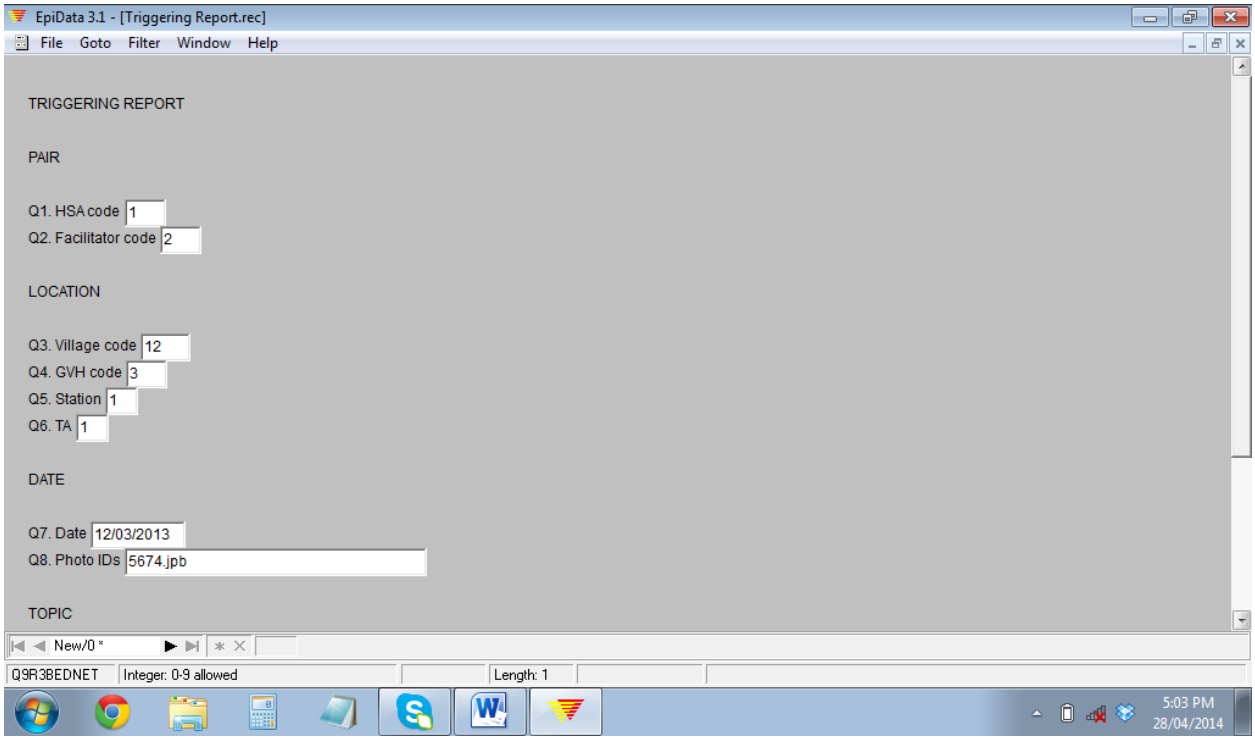
Step 1: Click on Enter Data



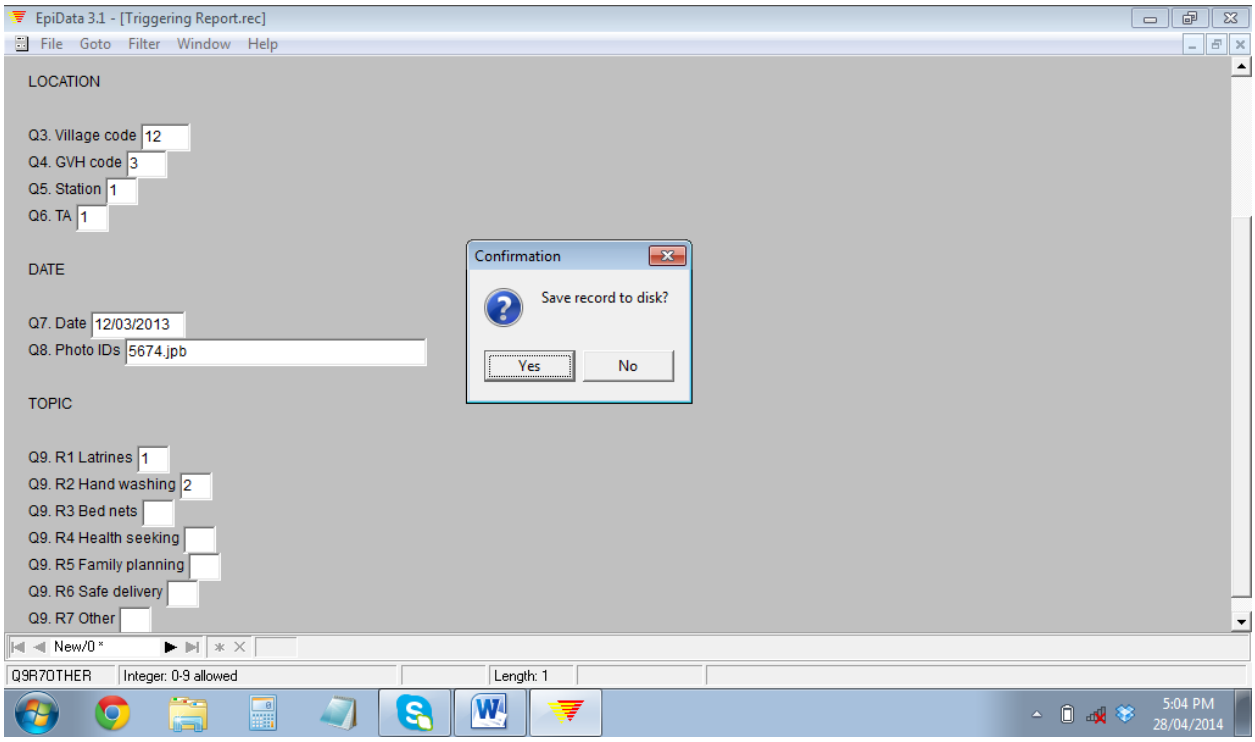
Step 2: Choose the file and click Open



Step 3: Enter the results for the first survey / report.



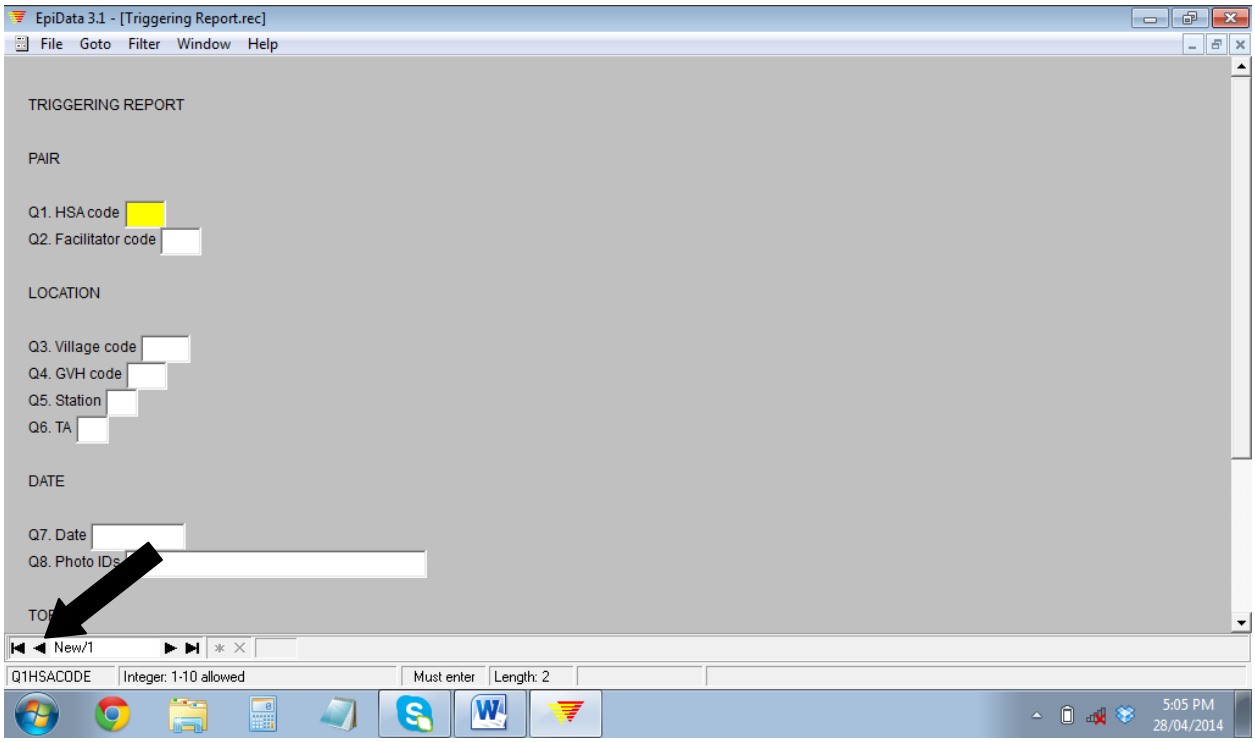
Step 4: When you reach the end of the form you will be asked if you want to save the record. Click Yes.



Step 5: Enter the next survey / report until all of them are complete.

HINT

Use the small arrows at the bottom of the screen to go back and forth between surveys you have already entered.



You don't need to enter the surveys in a particular order. If you miss one just enter it at any time. You will be able to re-order them later during the analysis.

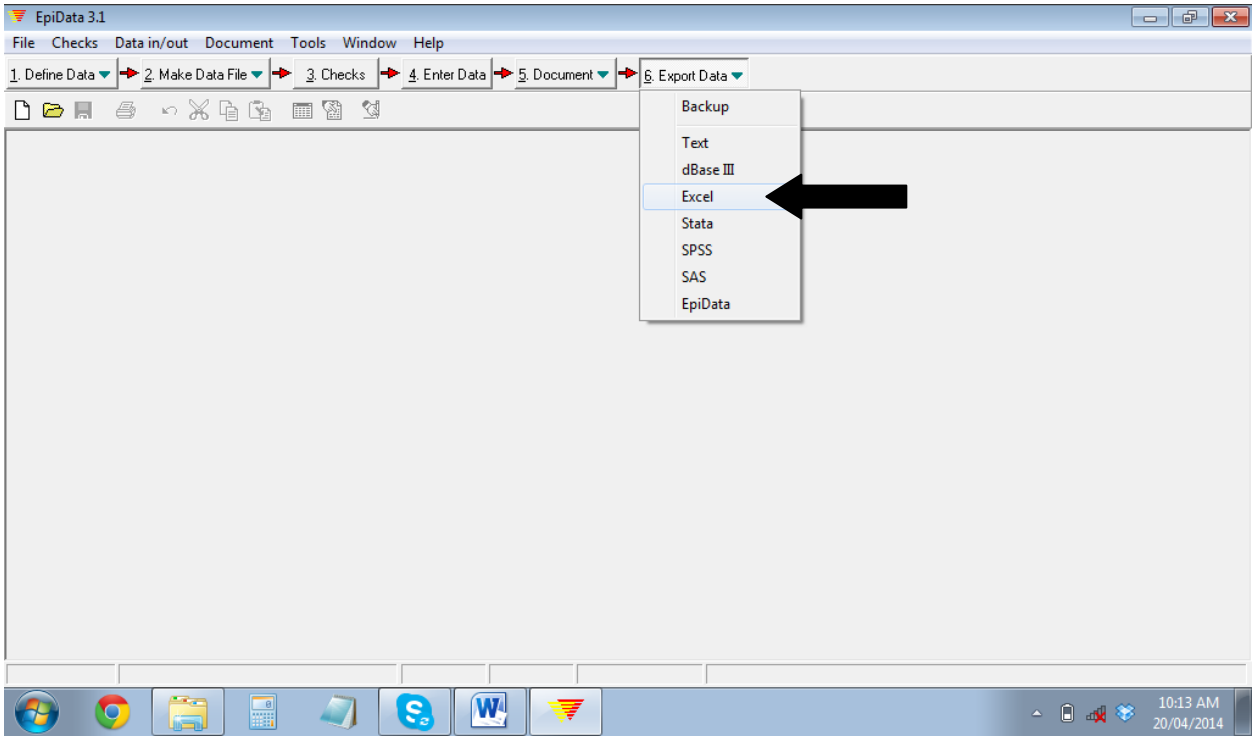
3.3 Exporting from EpiData

Once you have finished the data entry you need to export the data from EpiData to Excel. There are two ways to do this. The easiest way is to export directly to Excel. However, for surveys with a large number of questions this sometimes causes the questions at the end to be cut off.

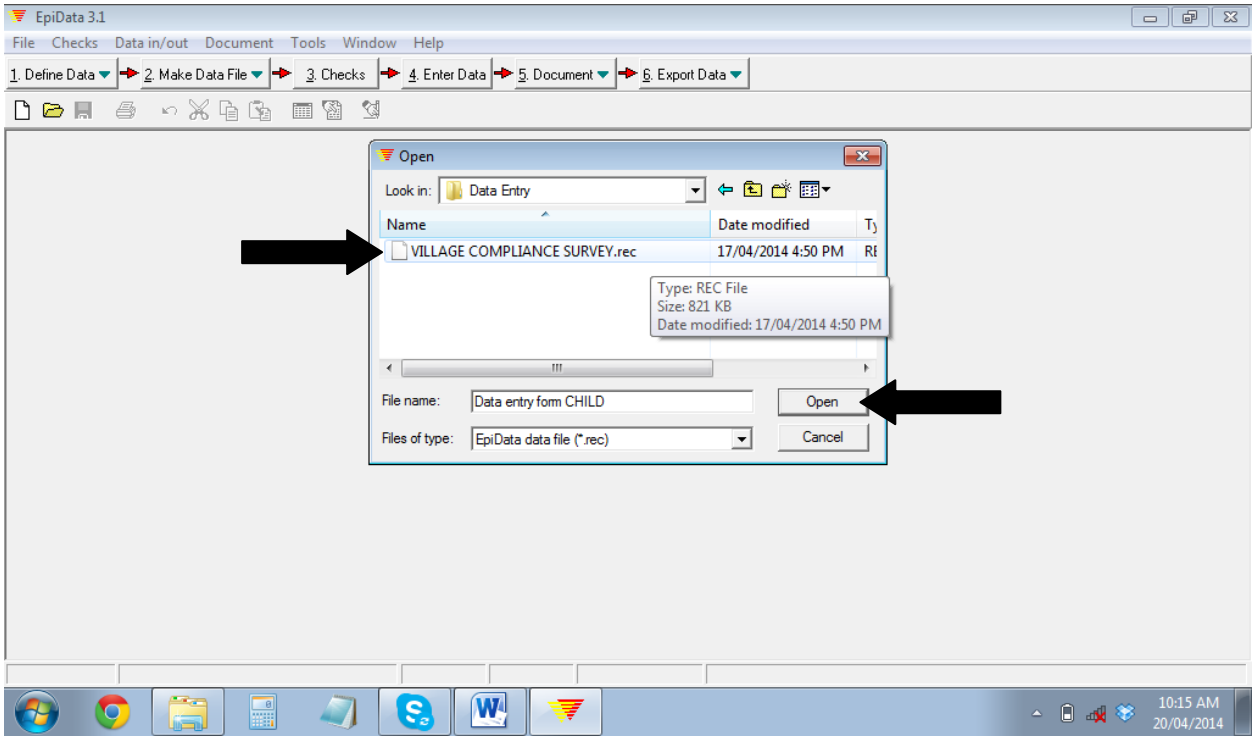
If you experience this problem then you should export from EpiData to a text file and then import the text file into Excel. Steps for both methods are shown below.

3.3.1 Exporting directly to Excel

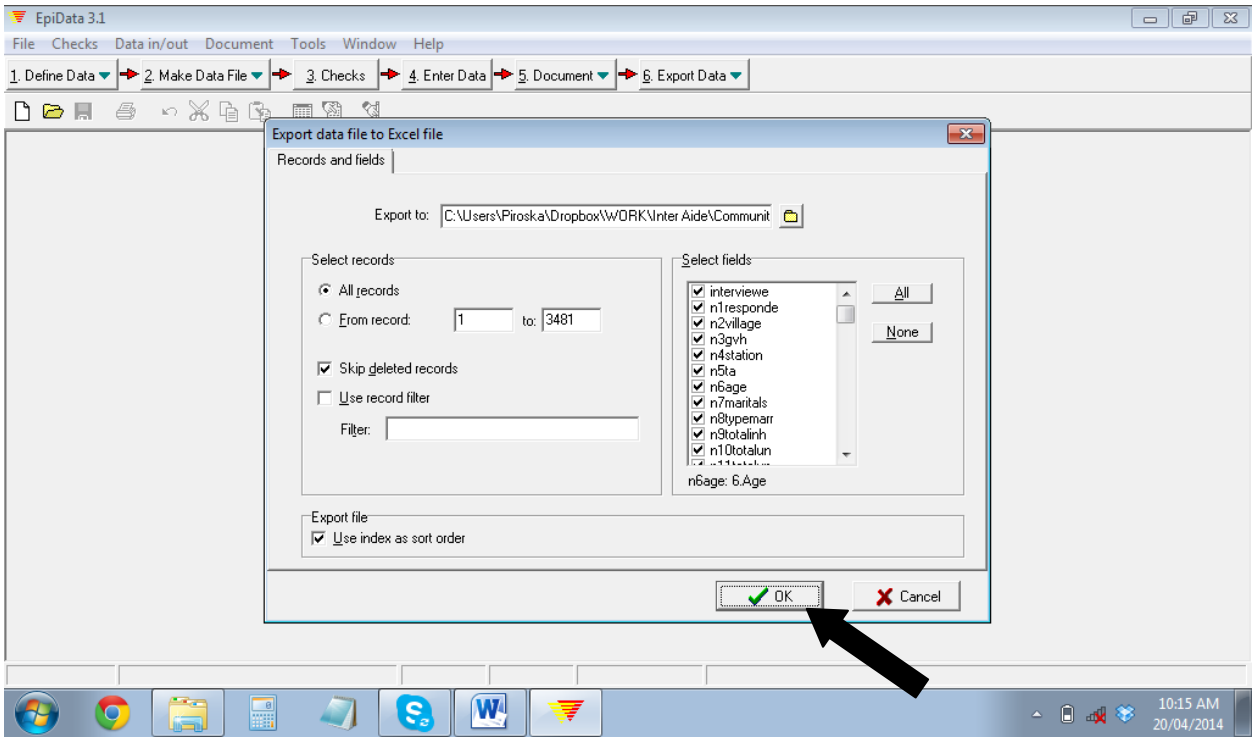
Step 1: Open EpiData. Click on the Export Data button and choose Excel.



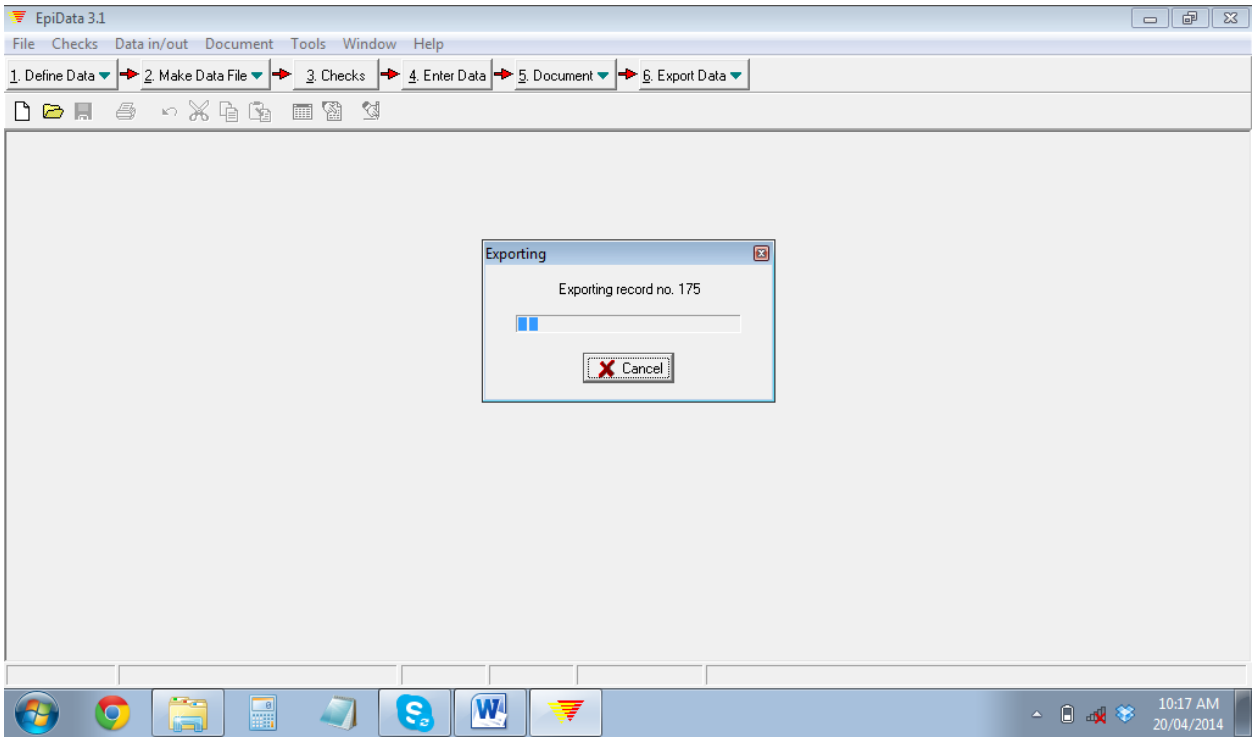
Step 2: Choose the EpiData file that you want to export and click Open.



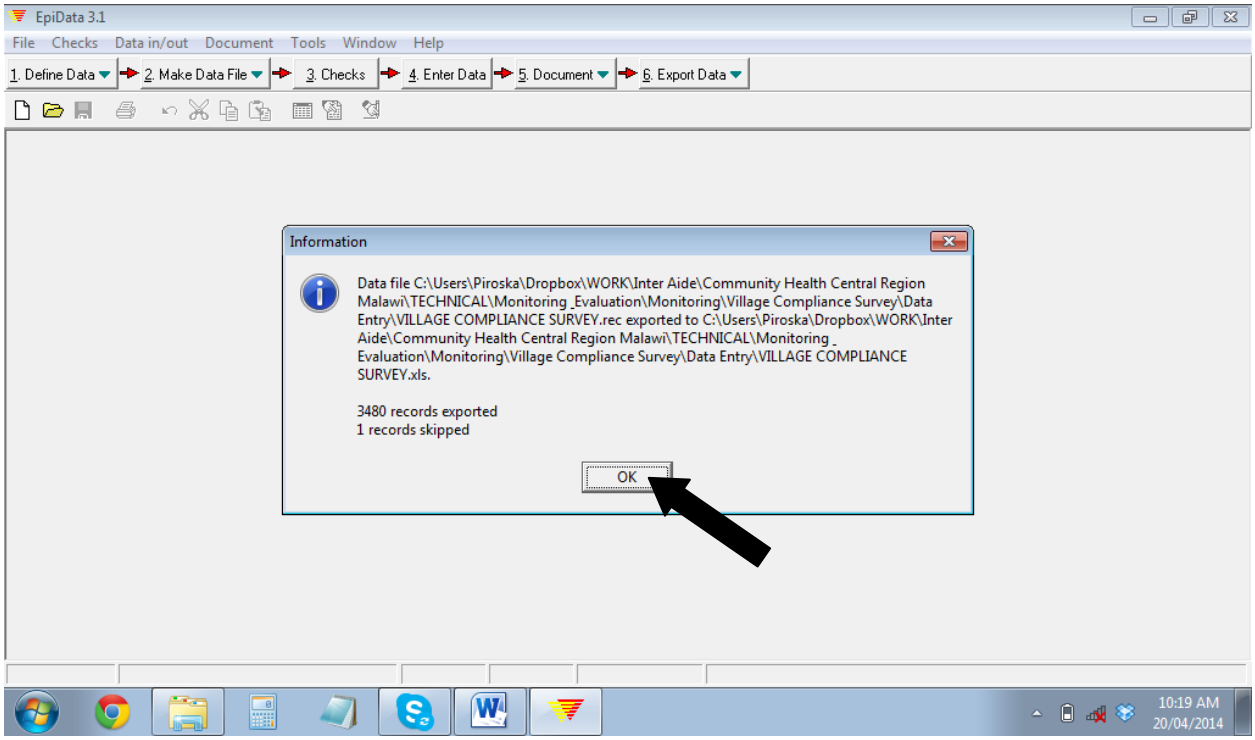
Step 3: The next screen allows you to choose which questions and records you want to export. Normally you don't need to change any of these settings. Just click OK.



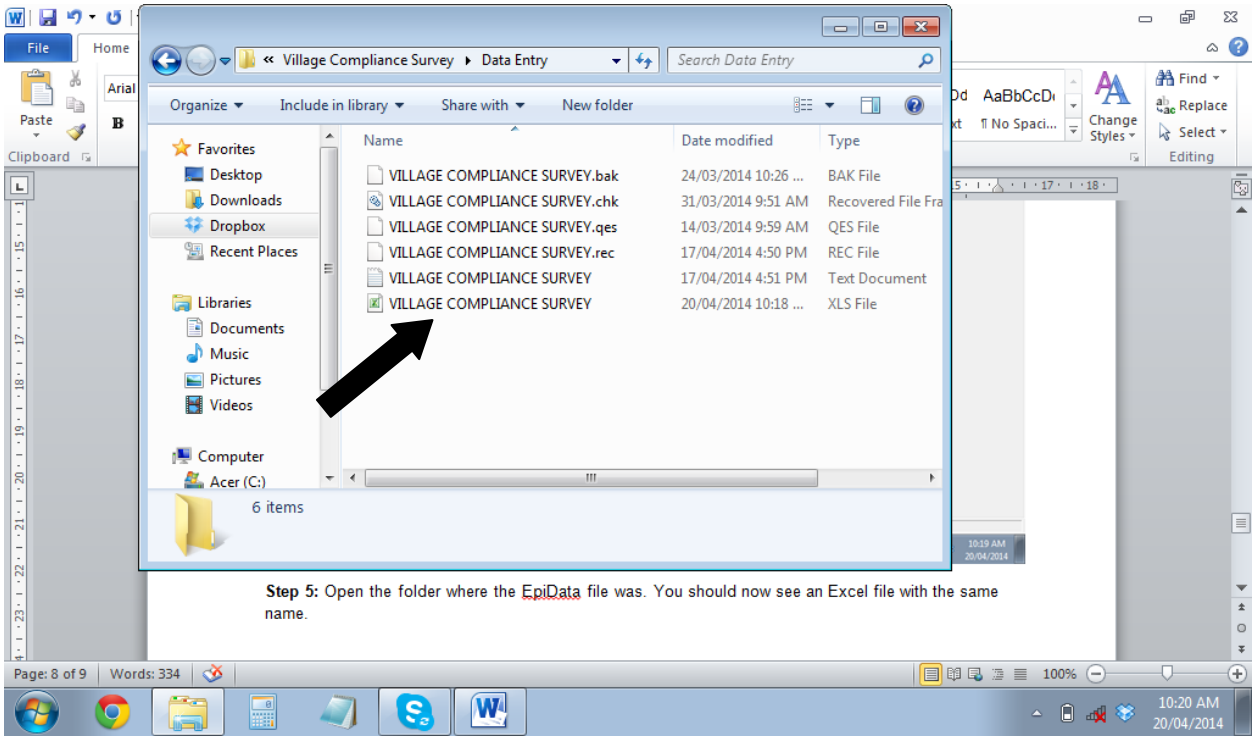
Step 4: You will see a progress bar showing the records being exported.



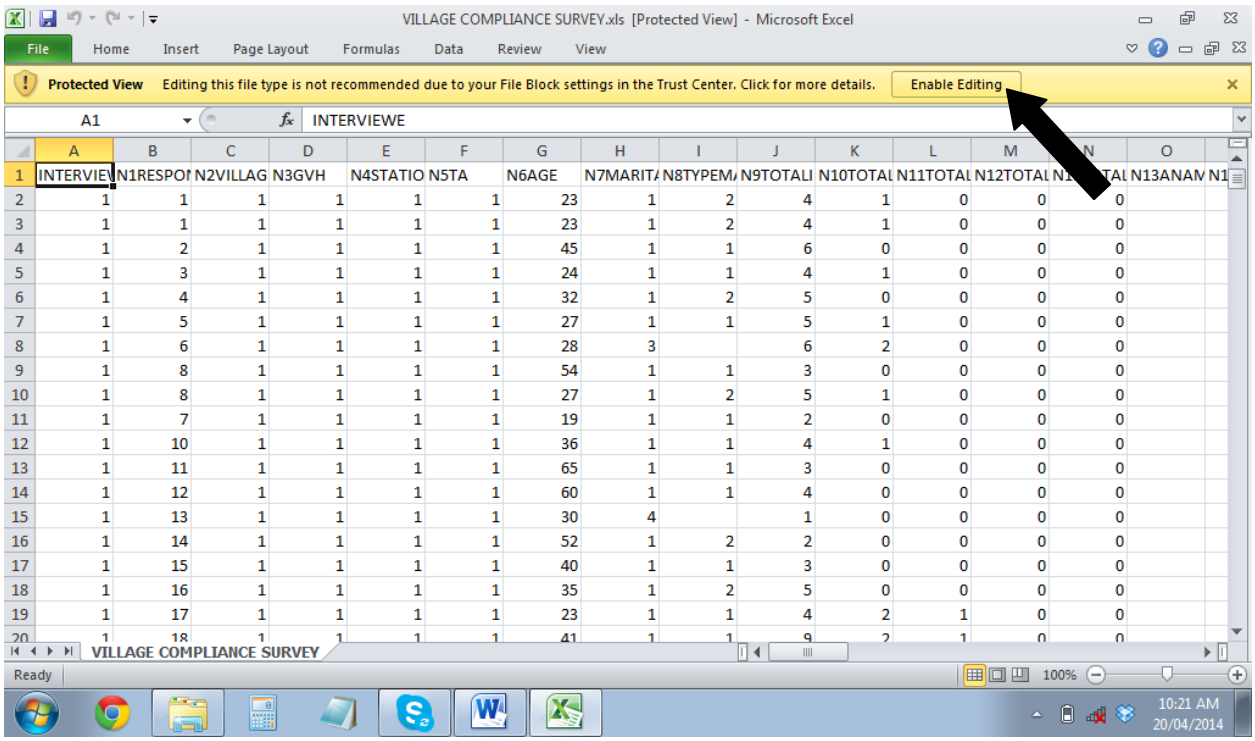
Step 4: When it has finished you will see a notice saying that the export has been completed. Click OK.



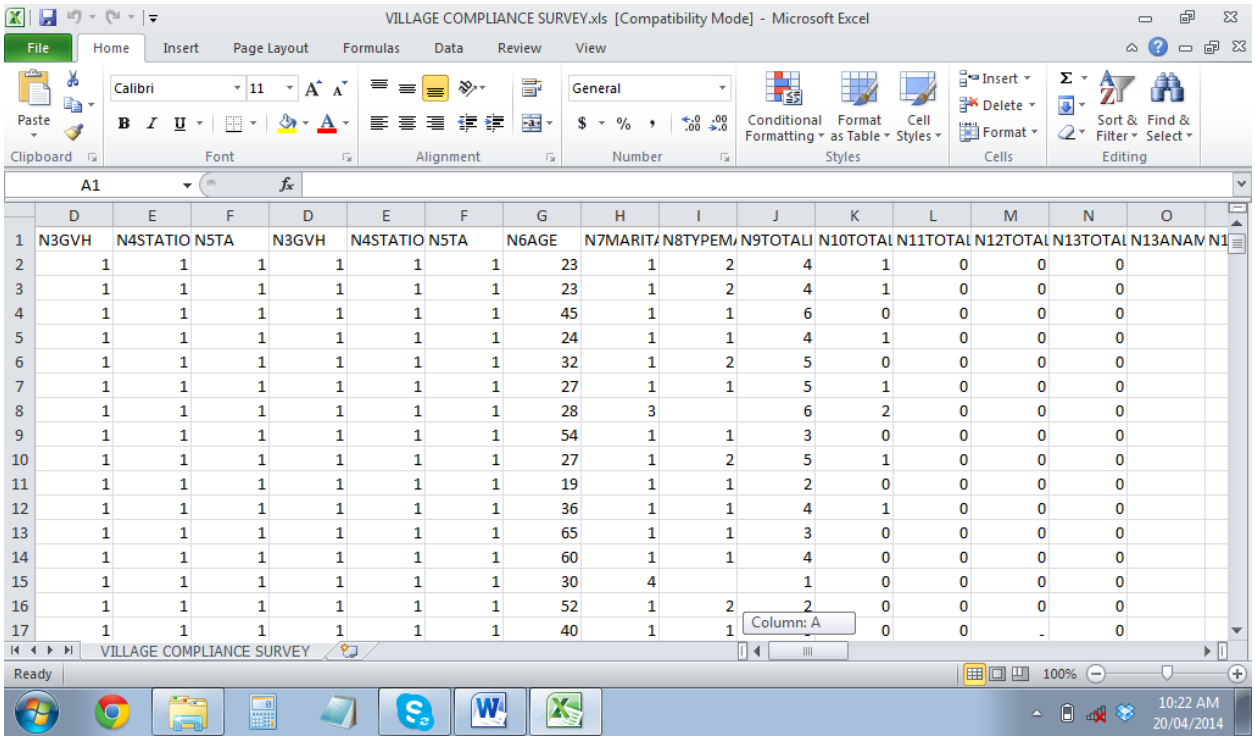
Step 5: Open the folder where the EpiData file was. You should now see an Excel file with the same name. Double click to open the Excel file.



Step 6: When the file opens you should see columns for each of the questions and rows for each of the surveys. If there is a yellow bar at the top saying Protected View click the Enable Editing button.



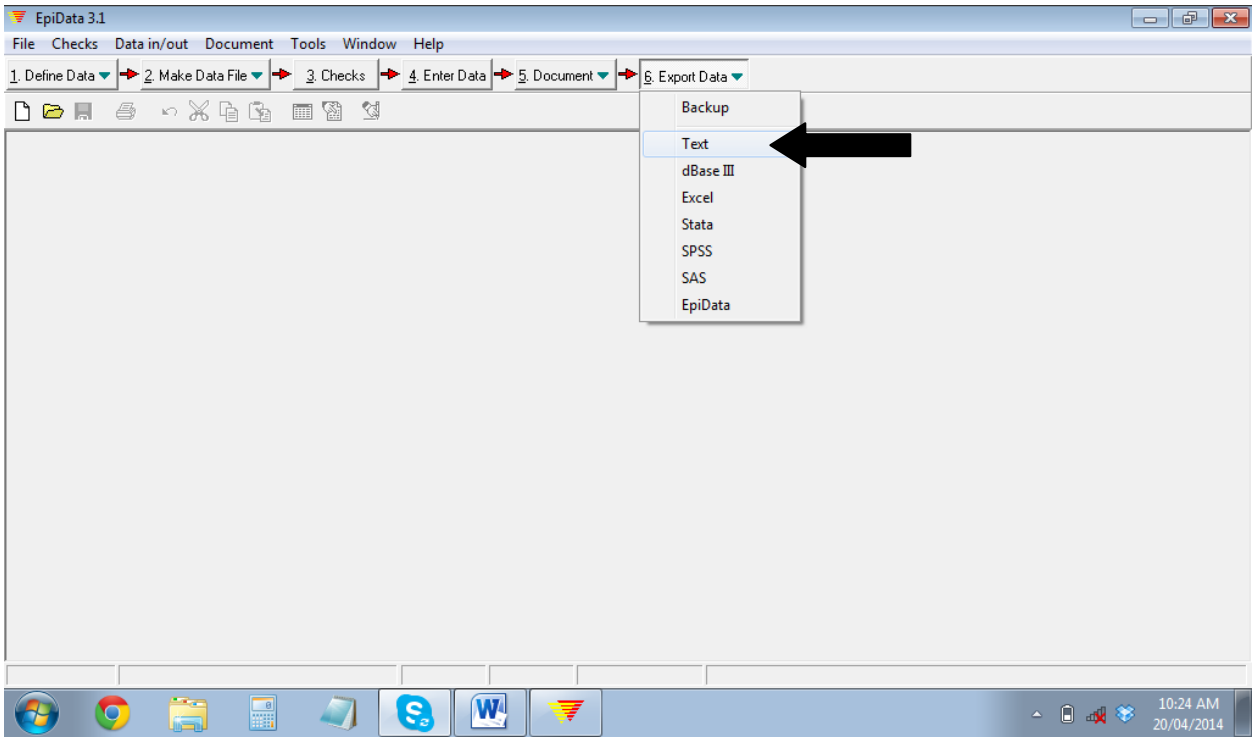
Step 7: The tools at the top of the Excel spreadsheet should appear. You are now ready for data cleaning.



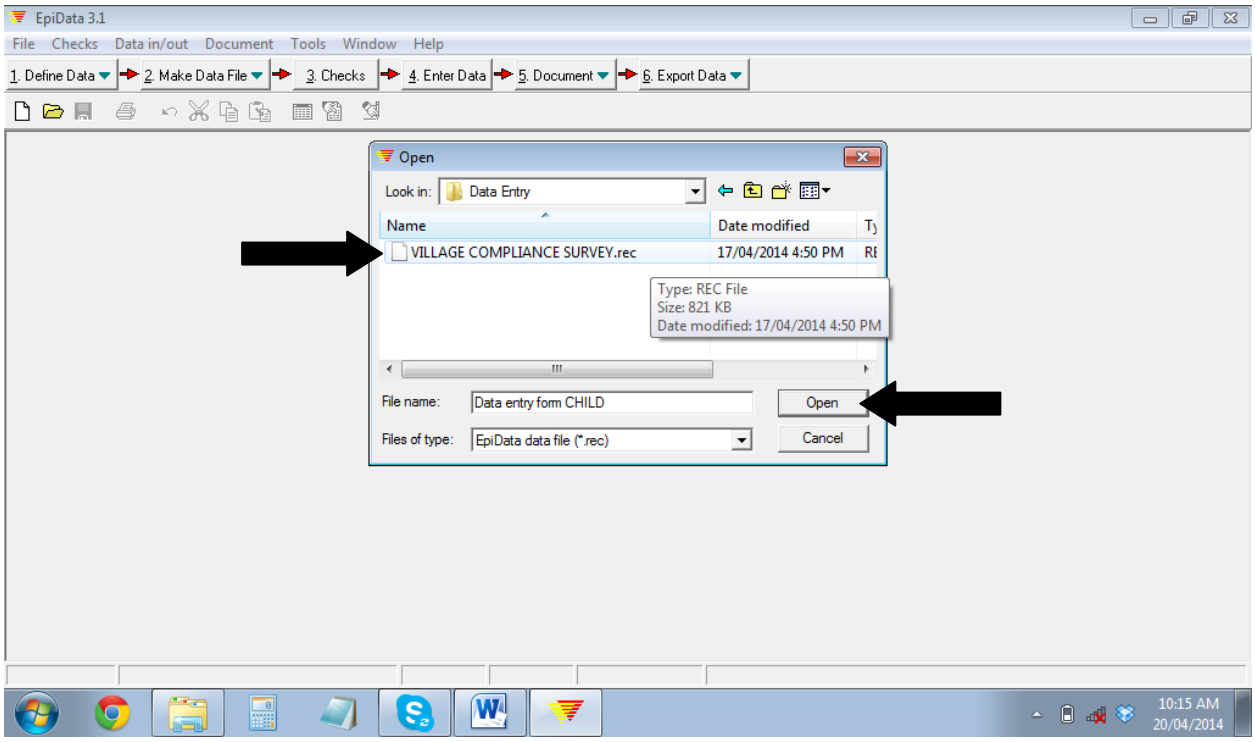
	D	E	F	D	E	F	G	H	I	J	K	L	M	N	O
1	N3GVH	N4STATIO	N5TA	N3GVH	N4STATIO	N5TA	N6AGE	N7MARITA	N8TYPEM/	N9TOTALI	N10TOTAL	N11TOTAL	N12TOTAL	N13TOTAL	N13ANAM
2	1	1	1	1	1	1	23	1	2	4	1	0	0	0	
3	1	1	1	1	1	1	23	1	2	4	1	0	0	0	
4	1	1	1	1	1	1	45	1	1	6	0	0	0	0	
5	1	1	1	1	1	1	24	1	1	4	1	0	0	0	
6	1	1	1	1	1	1	32	1	2	5	0	0	0	0	
7	1	1	1	1	1	1	27	1	1	5	1	0	0	0	
8	1	1	1	1	1	1	28	3		6	2	0	0	0	
9	1	1	1	1	1	1	54	1	1	3	0	0	0	0	
10	1	1	1	1	1	1	27	1	2	5	1	0	0	0	
11	1	1	1	1	1	1	19	1	1	2	0	0	0	0	
12	1	1	1	1	1	1	36	1	1	4	1	0	0	0	
13	1	1	1	1	1	1	65	1	1	3	0	0	0	0	
14	1	1	1	1	1	1	60	1	1	4	0	0	0	0	
15	1	1	1	1	1	1	30	4		1	0	0	0	0	
16	1	1	1	1	1	1	52	1	2	2	0	0	0	0	
17	1	1	1	1	1	1	40	1	1		0	0	-	0	

3.3.2 Exporting to a text file then importing into Excel

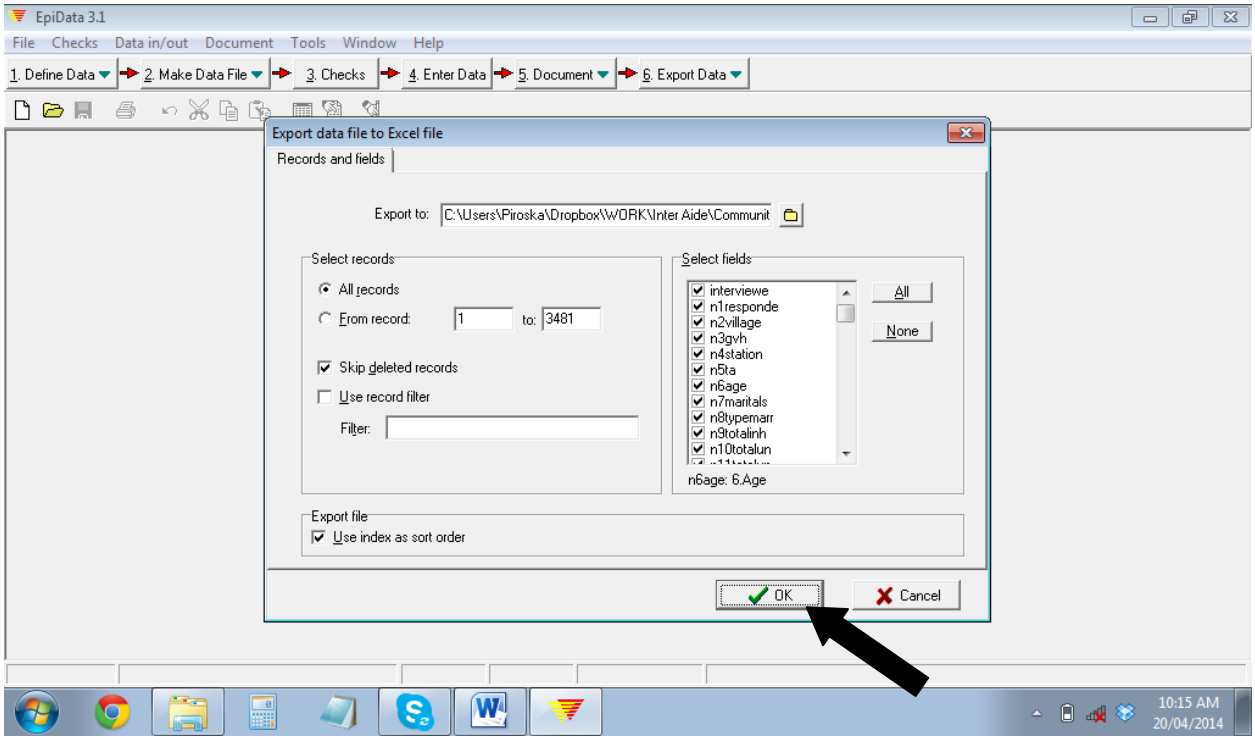
Step 1: Open EpiData. Click on the Export Data button and choose Text.



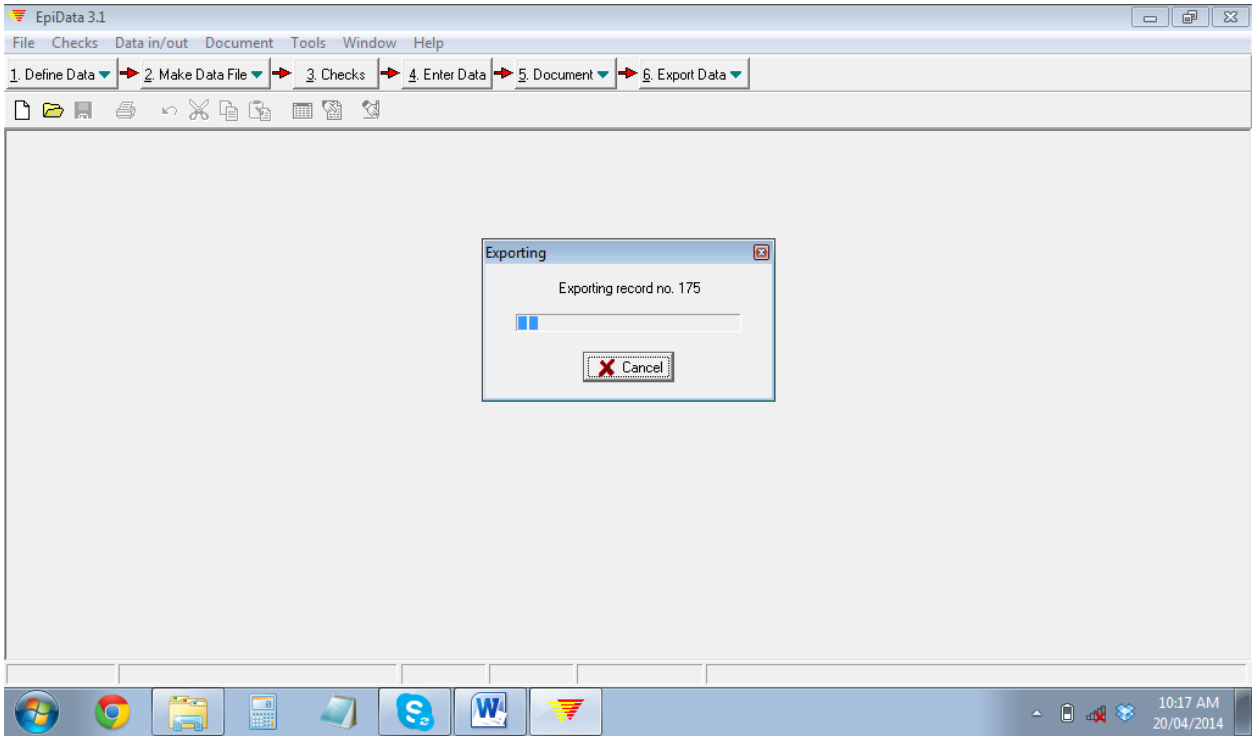
Step 2: Choose the EpiData file that you want to export and click Open.



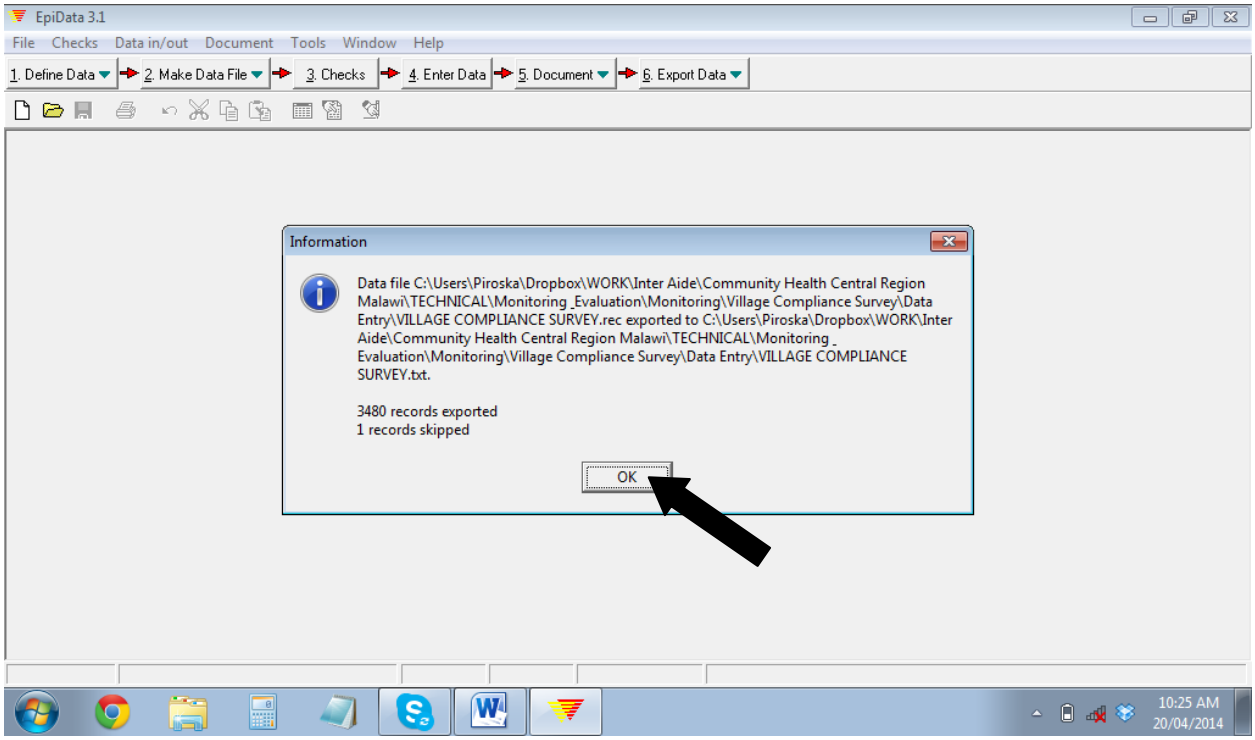
Step 3: The next screen allows you to choose which questions and records you want to export. Normally you don't need to change any of these settings. Just click OK.



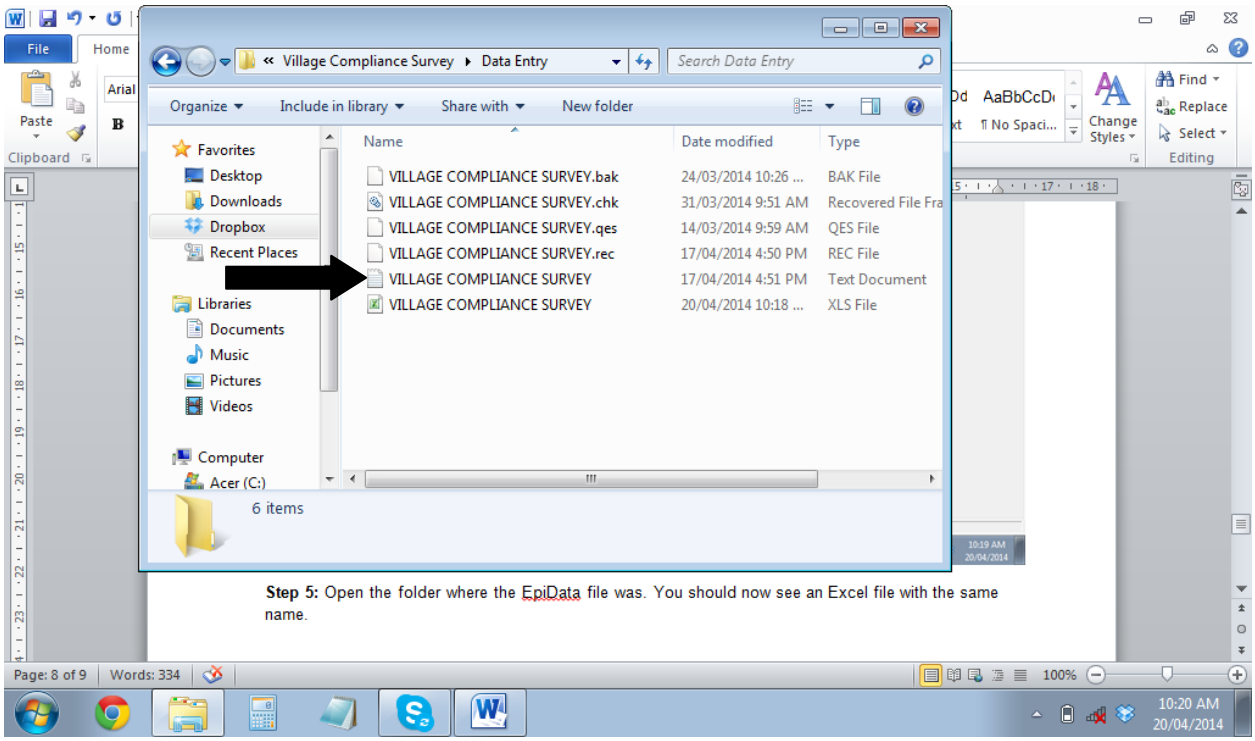
Step 4: You will see a progress bar showing the records being exported.



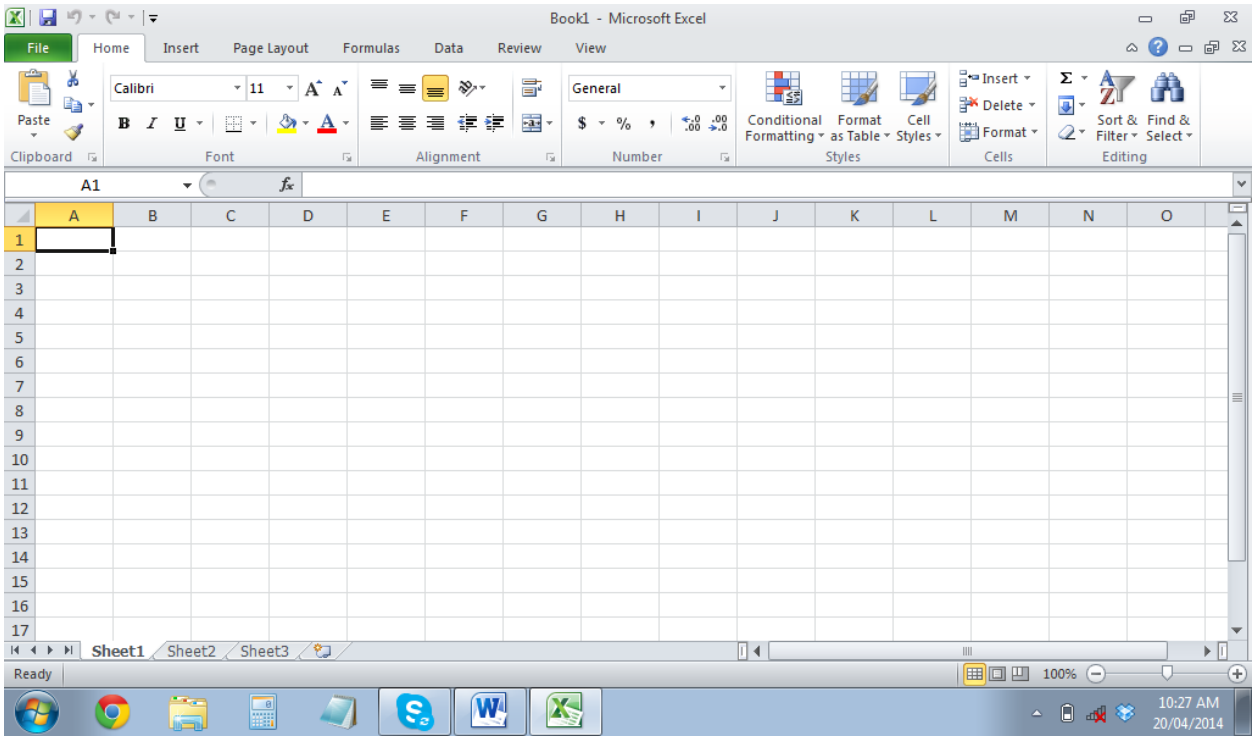
Step 4: When it has finished you will see a notice saying that the export has been completed. Click OK.



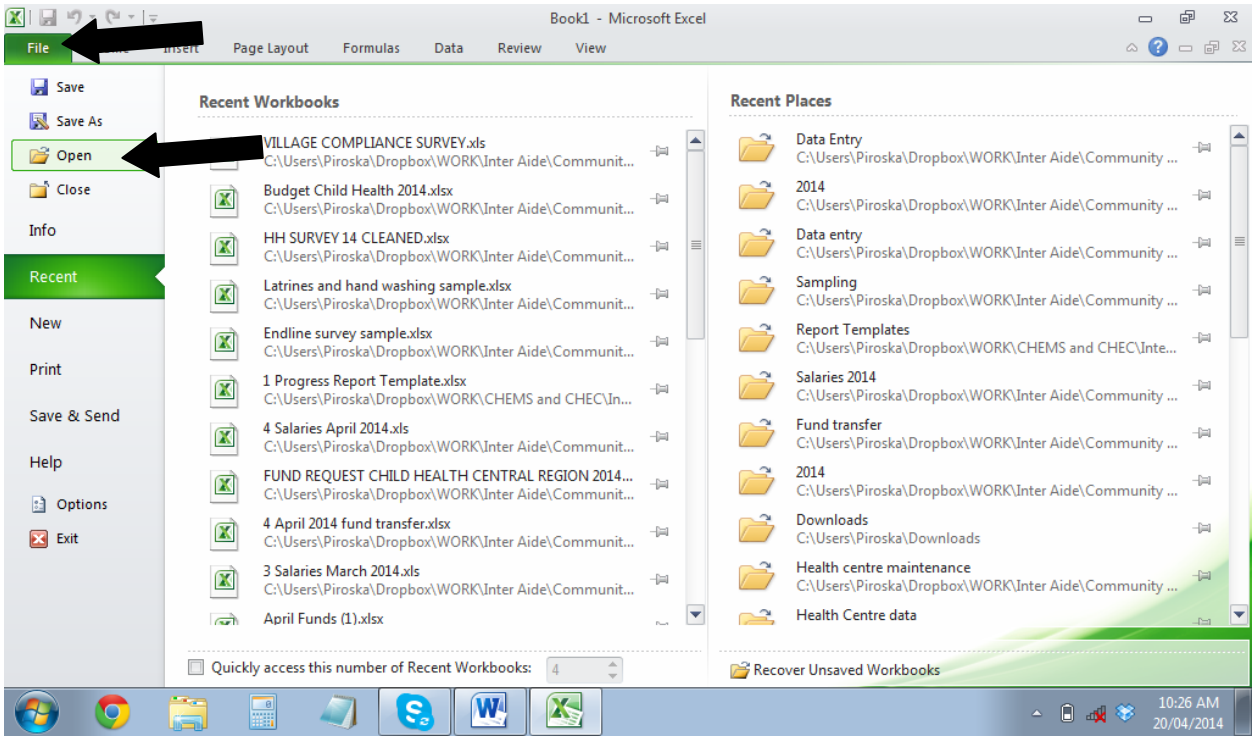
Step 5: Open the folder where the EpiData file was. You should now see a Text file with the same name. There is no need to open it.



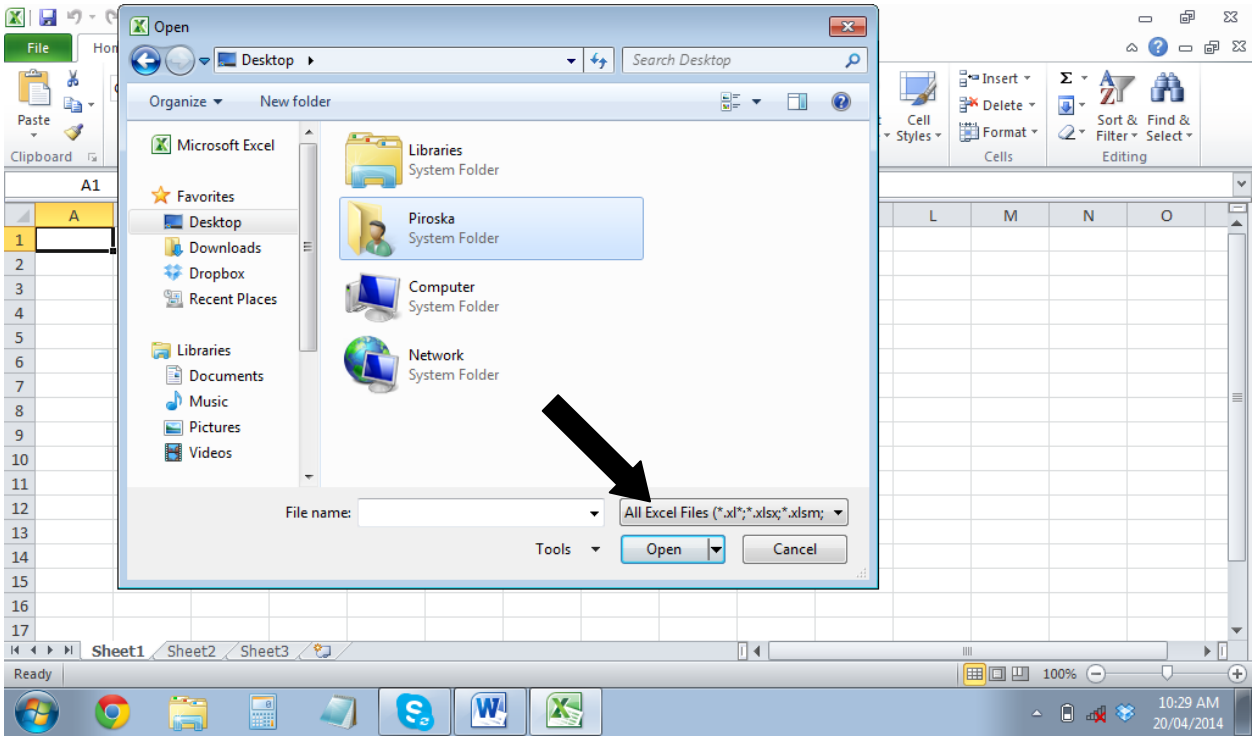
Step 6: Open a new blank spreadsheet in Excel.

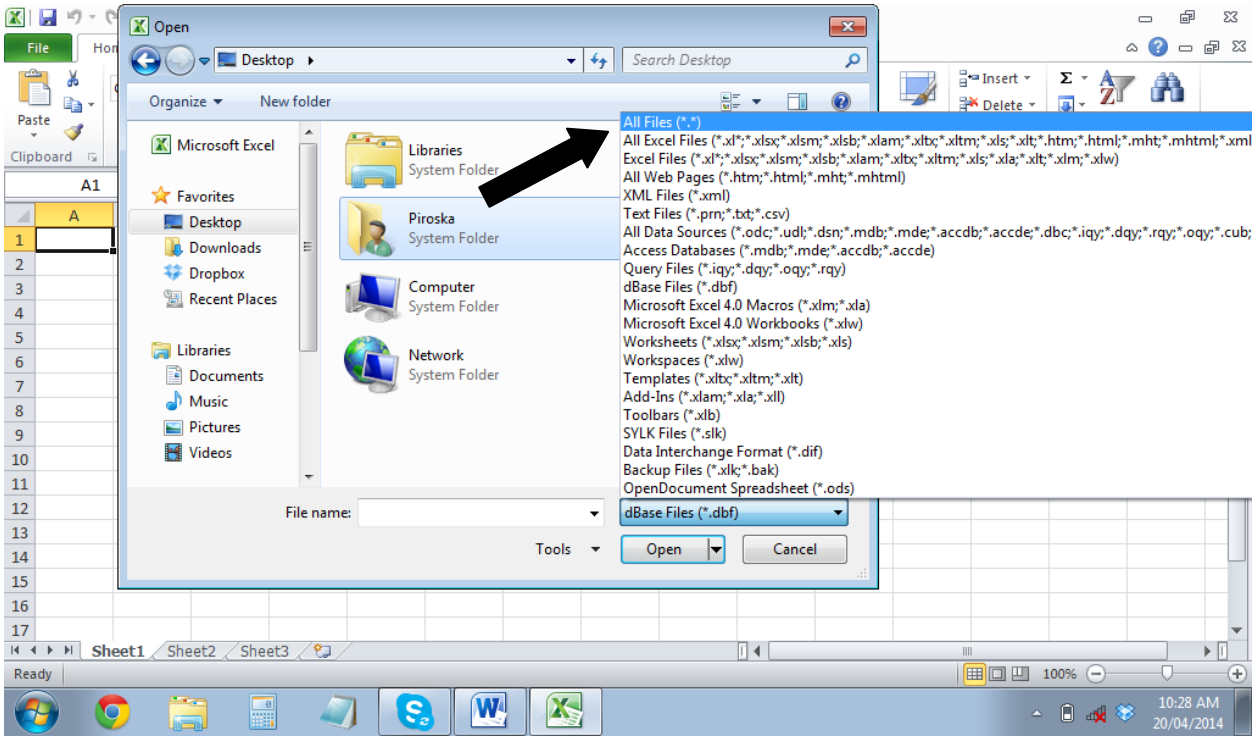


Step 7: Click on File, then Open.

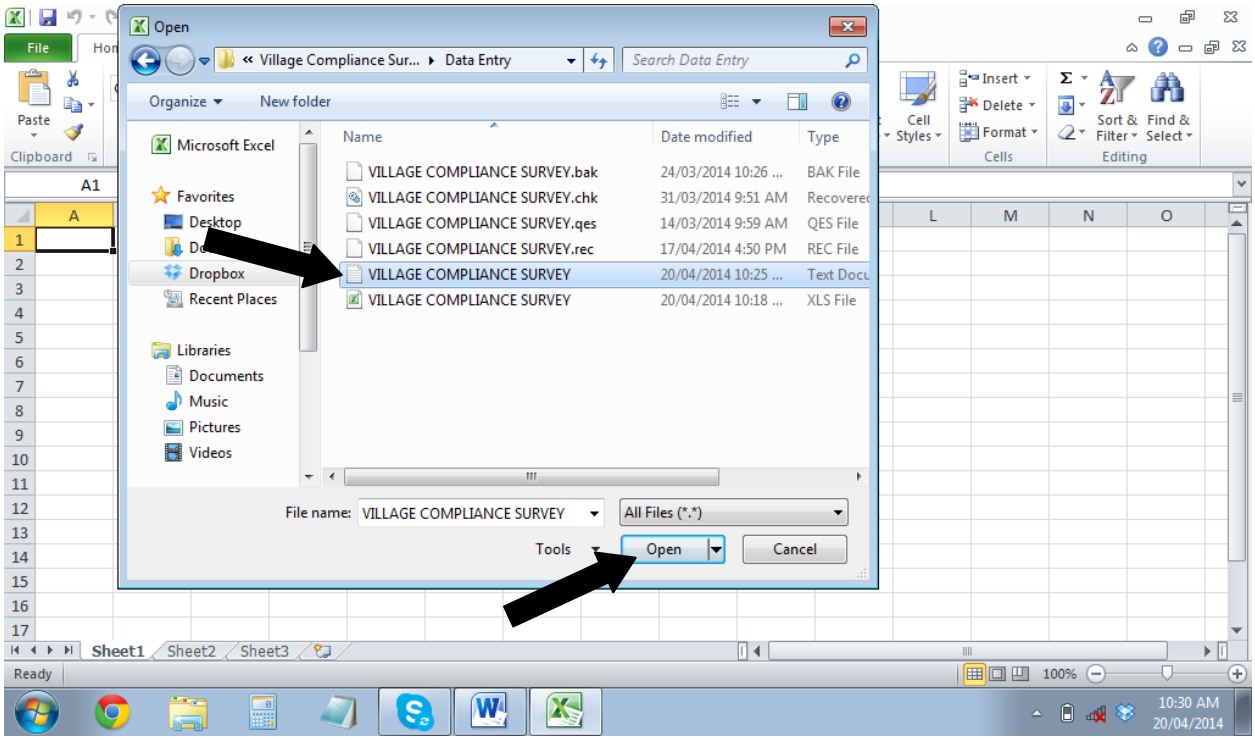


Step 8: Click on the button that says All Excel Files and change it to All Files.

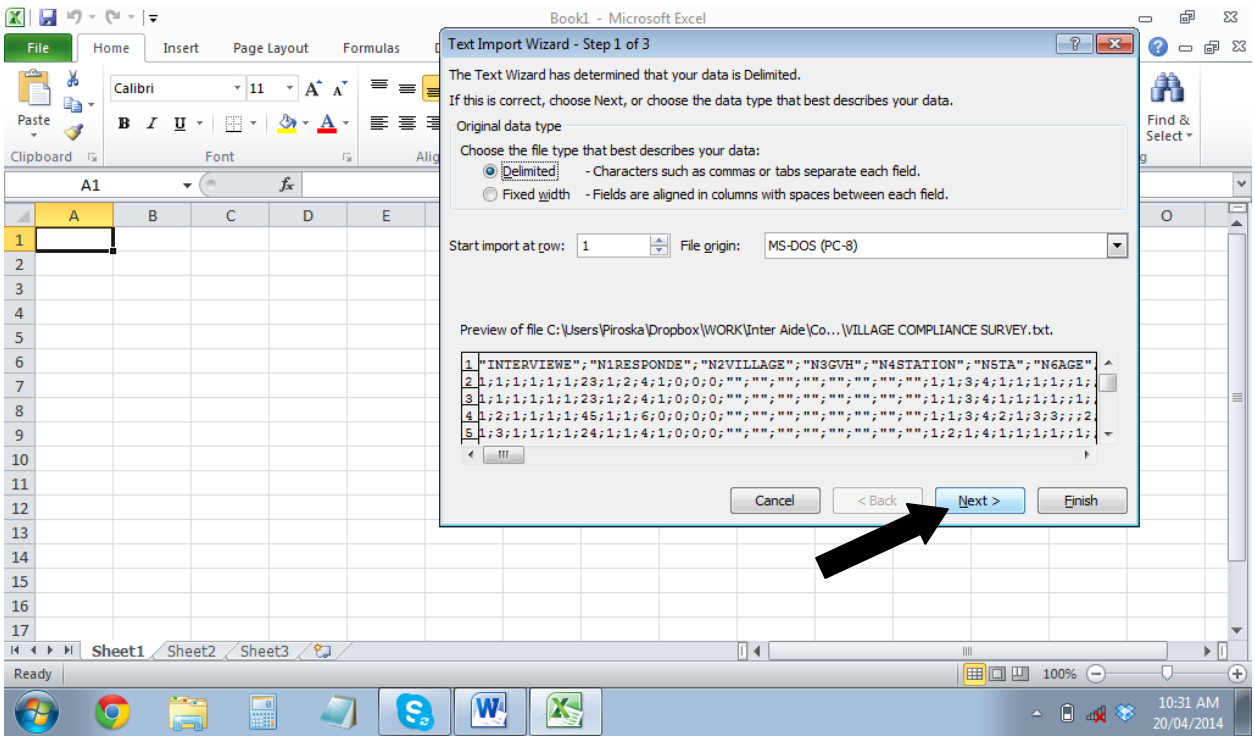




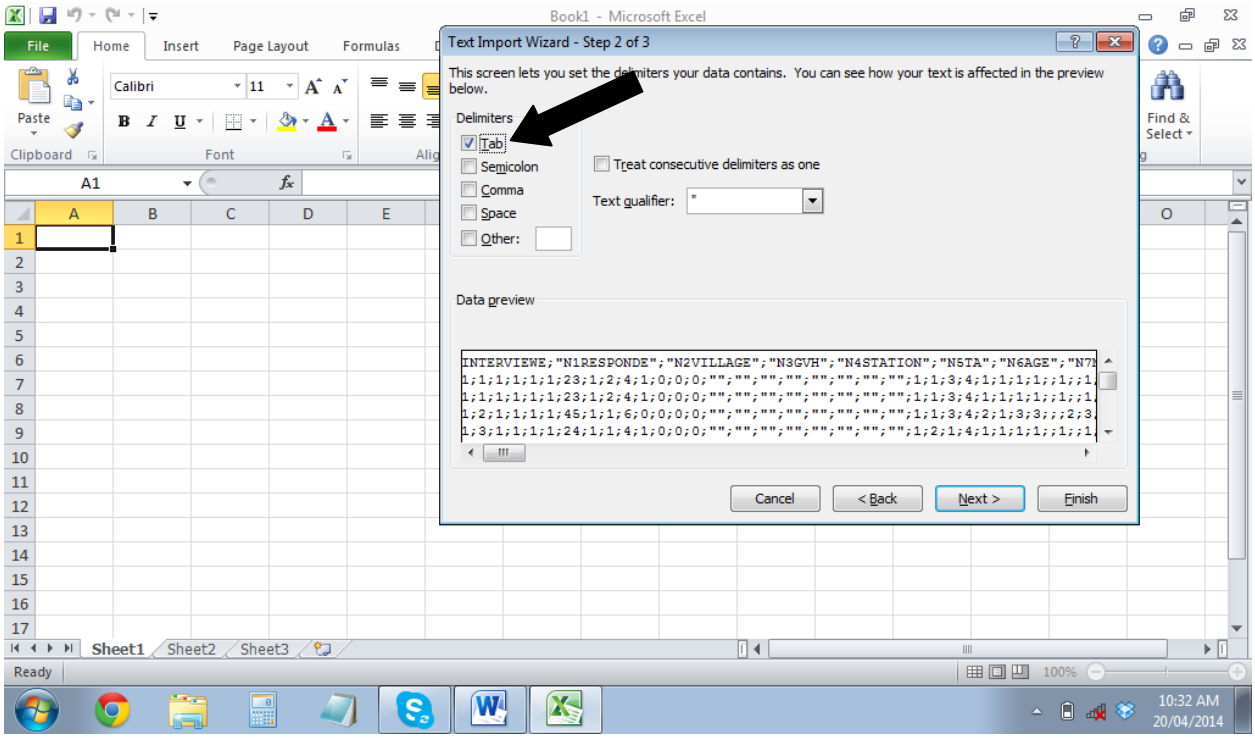
Step 8: Find the folder where the EpiData file was saved. You should see a Text file with the same name. Select it and then click Open.

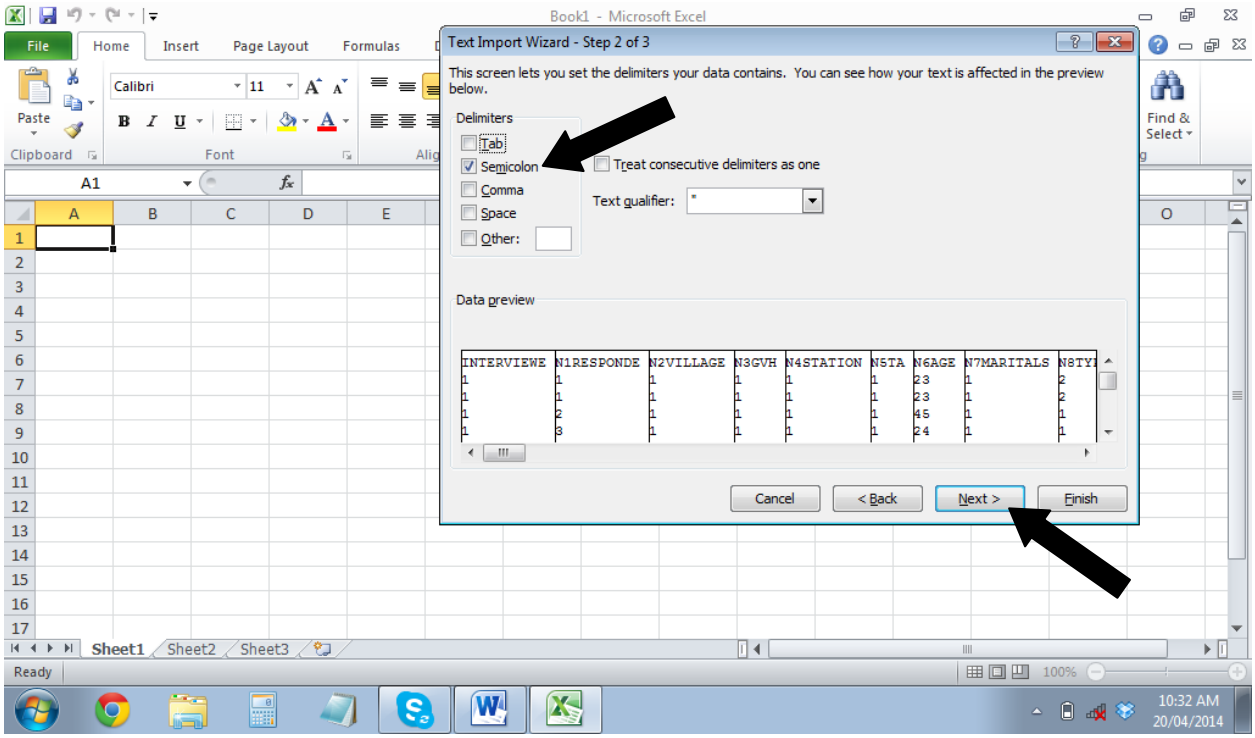


Step 9: A new window will open with some options. Leave the options as they are and just click Next.

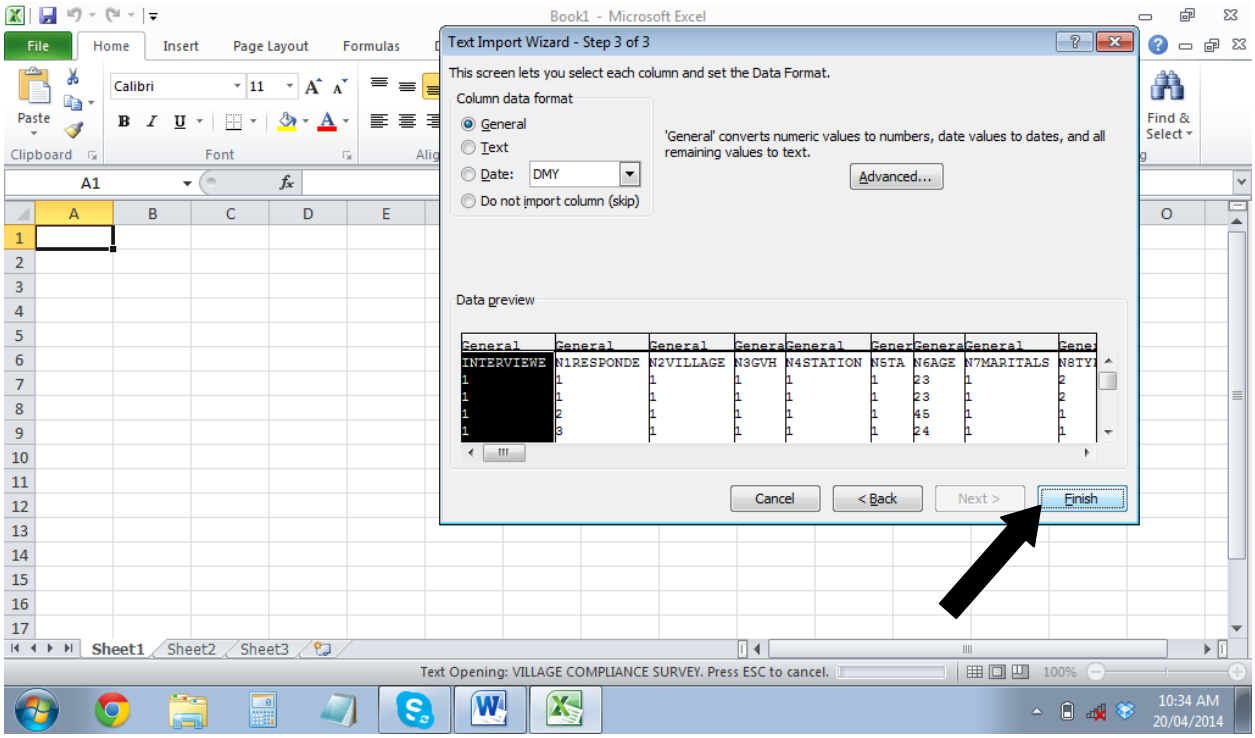


Step 10: On the next set of options un-tick Tab, then tick Semicolon. When you do this the data in the small window should appear correctly in columns. Finally, click Next.

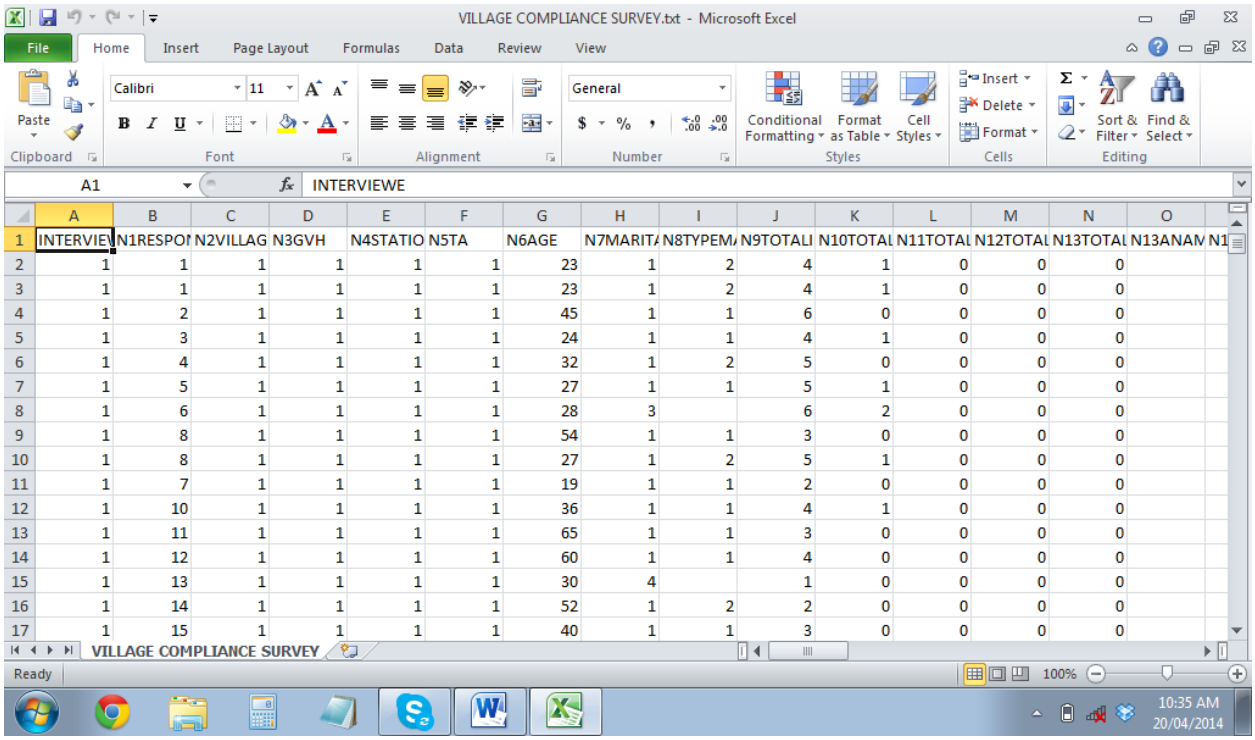




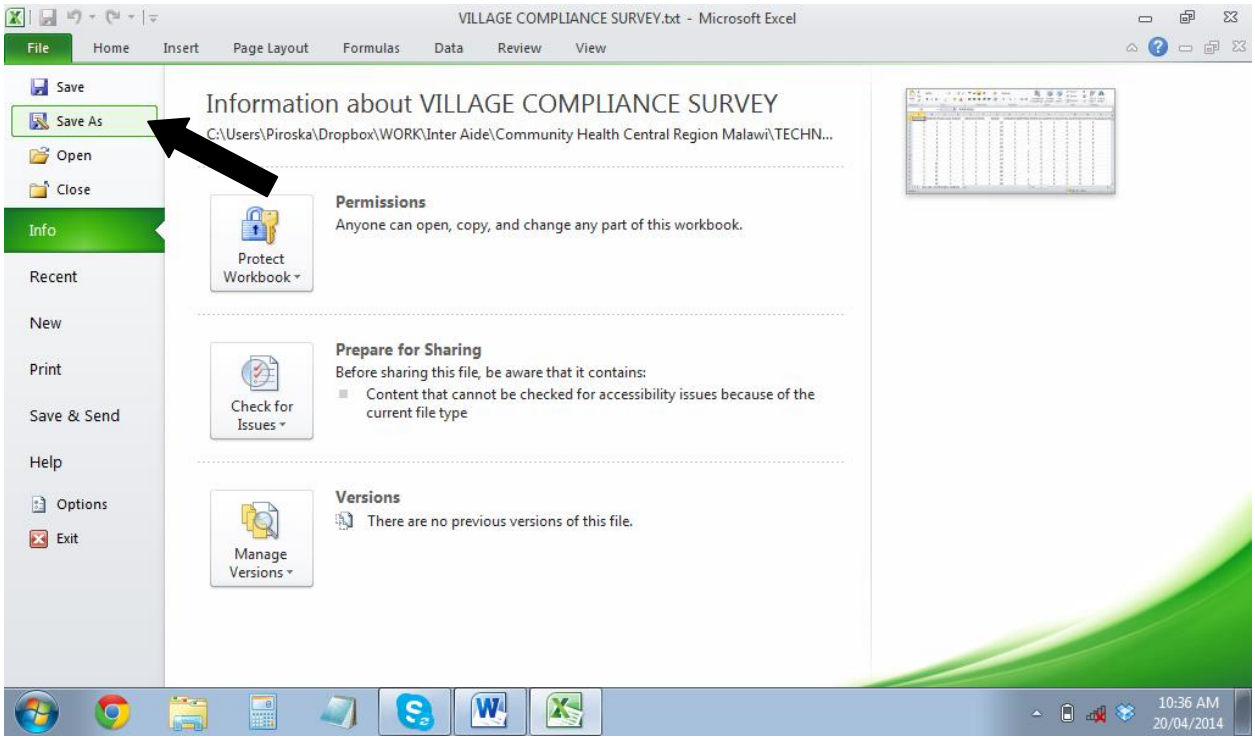
Step 11: On the final screen click Finish.



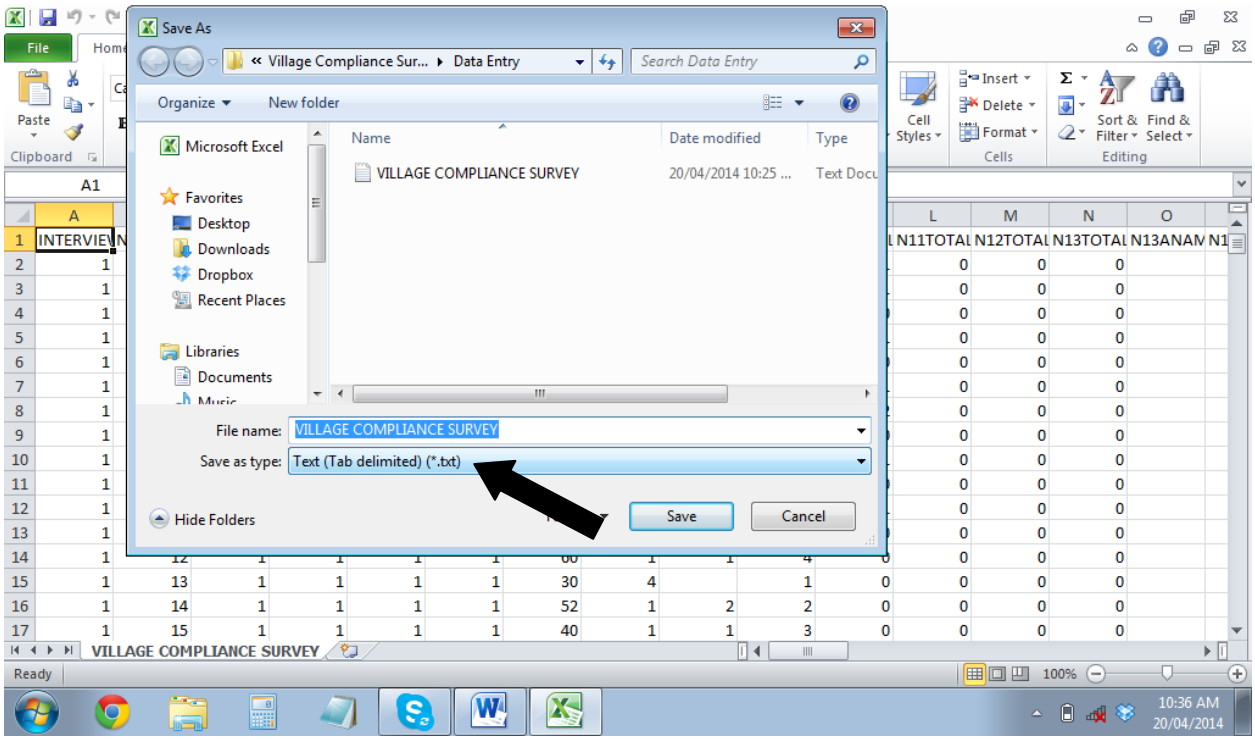
Step 11: On the final screen click Finish. Your data should now appear in Excel.

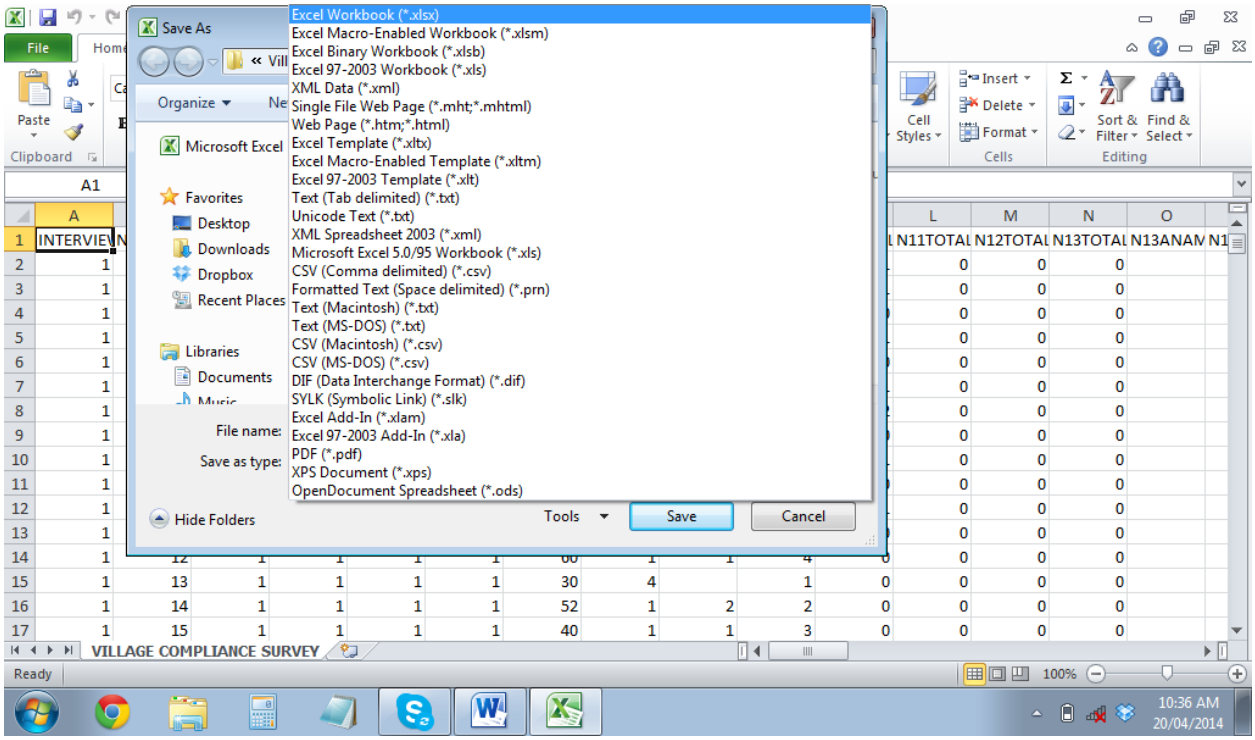


Step 12: To save it as an Excel file go to File, then choose Save As.

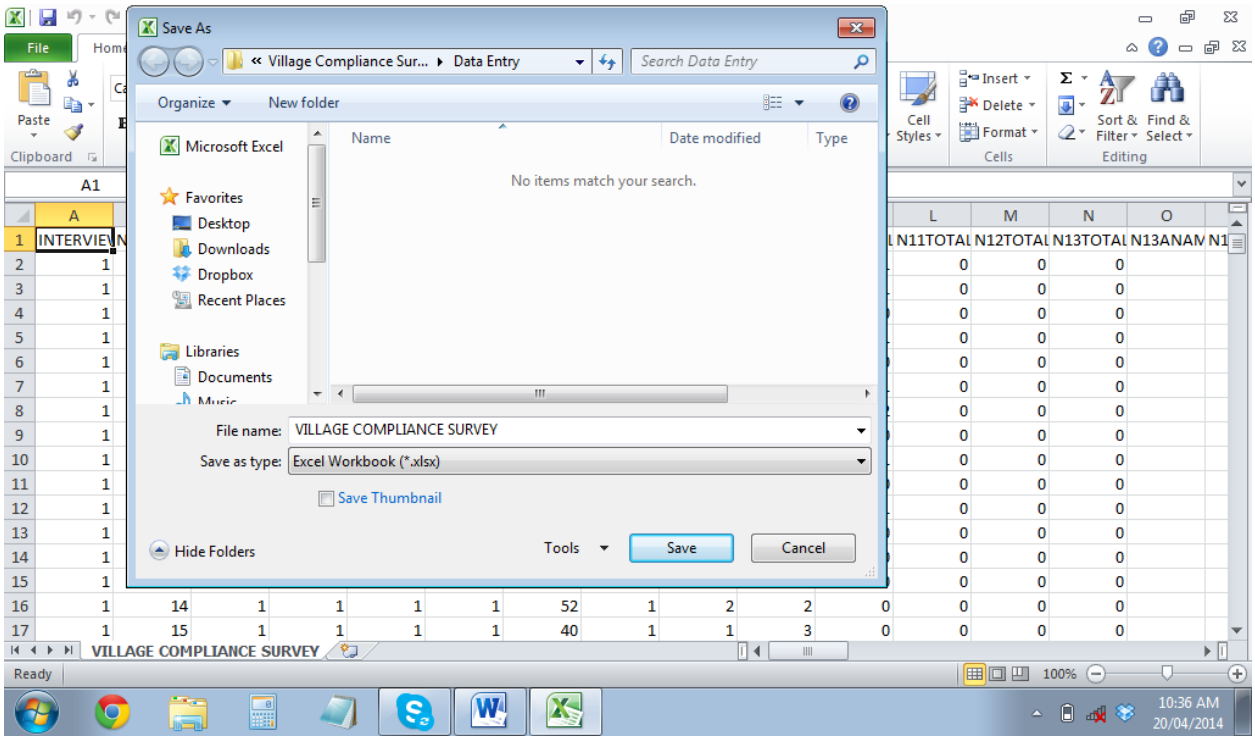


Step 13: Click the button next to Save As Type and choose Excel Workbook from the list.





Step 14: Click Save. You are now ready for data cleaning.

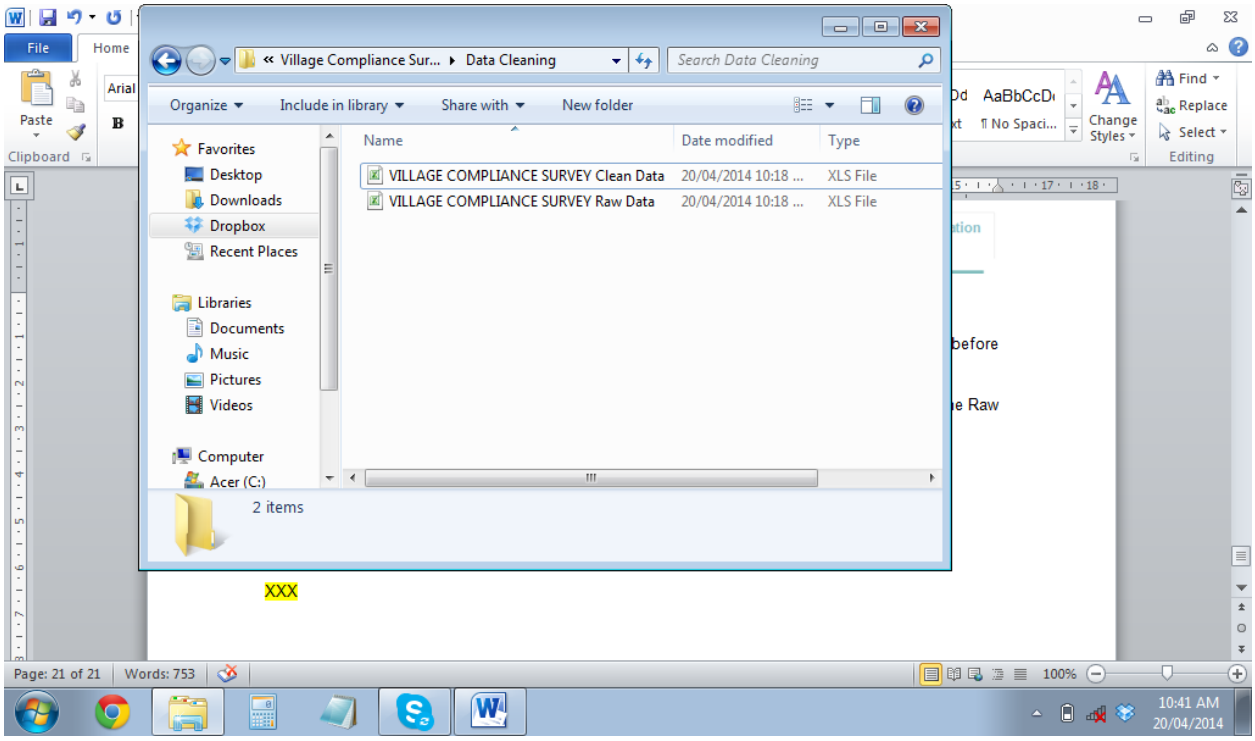


4 Data Cleaning

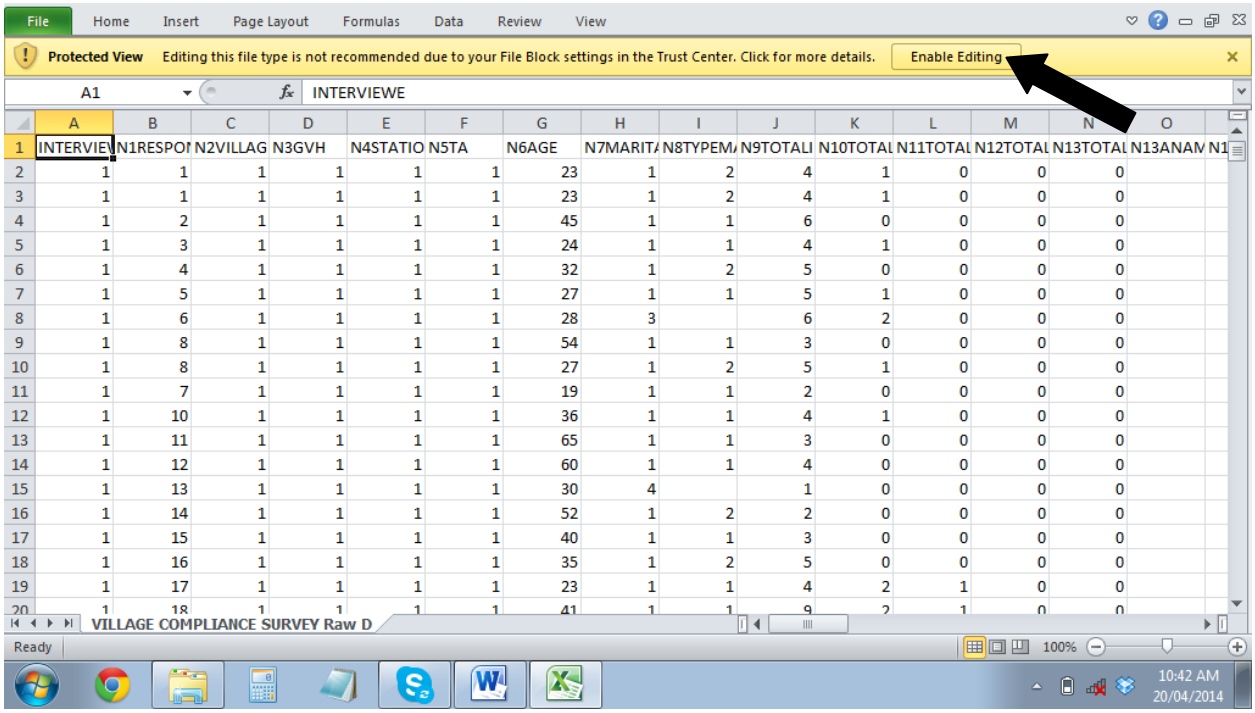
Data cleaning is the process of checking data for obvious errors and making corrections before analysis.

4.1 Setting up a filter

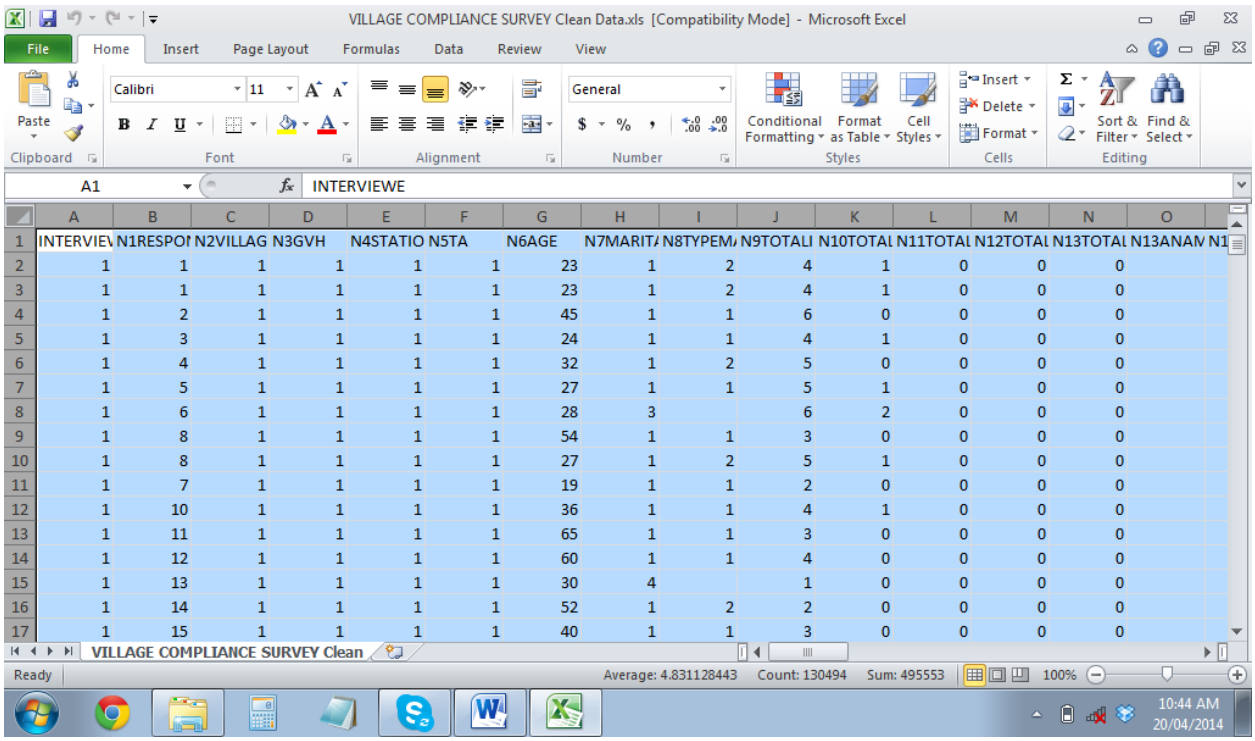
Step 1: Make two copies of your Excel file. Label one Raw Data and one Clean Data. Keep the Raw Data as a backup. Only work on the Clean Data version.



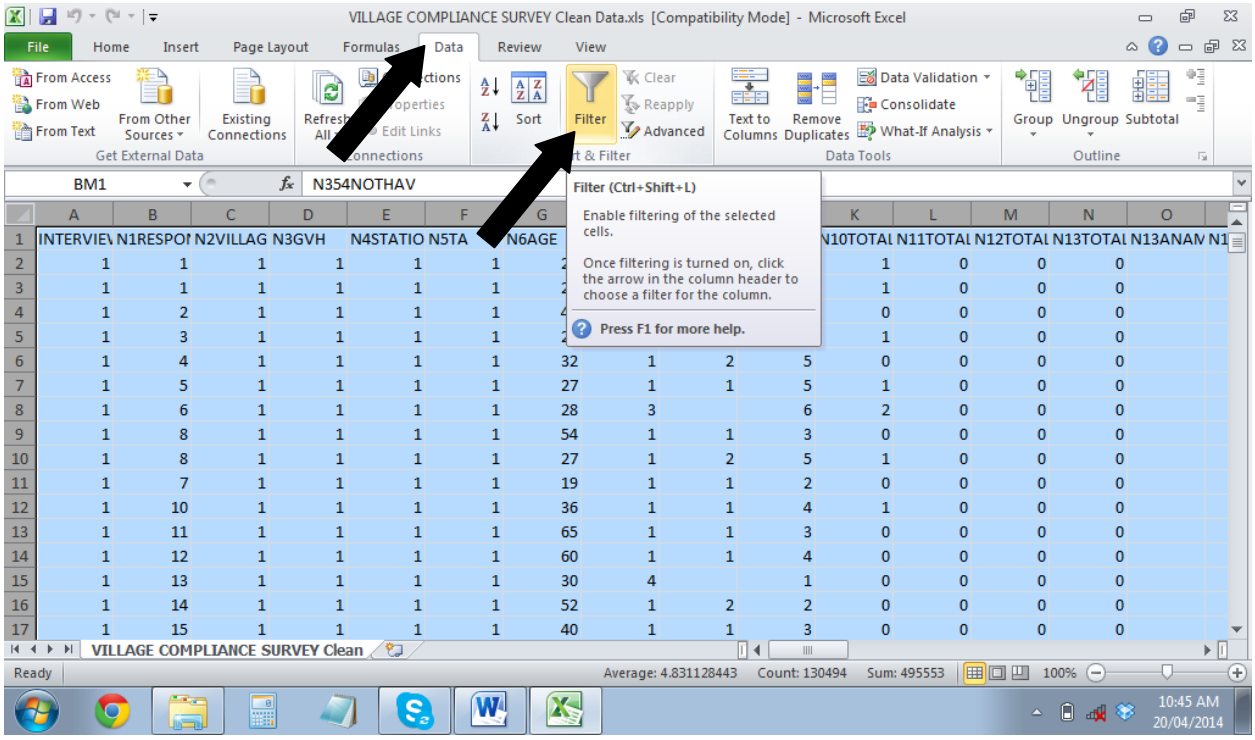
Step 2: Open the Clean Data file in Excel. If there is a yellow bar at the top saying Protected View click the Enable Editing button.



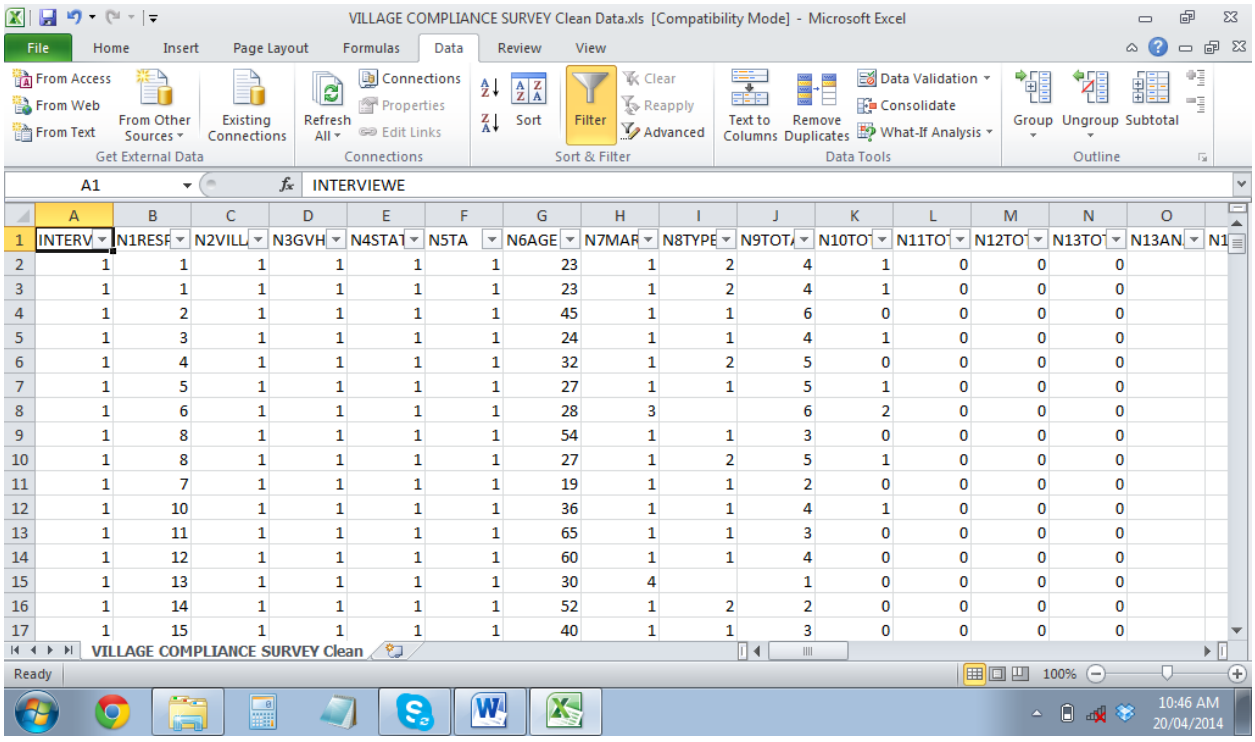
Step 3: Select all the rows and columns in the spreadsheet by clicking in the very top left corner. Or press the Ctrl + A keys on the keyboard.



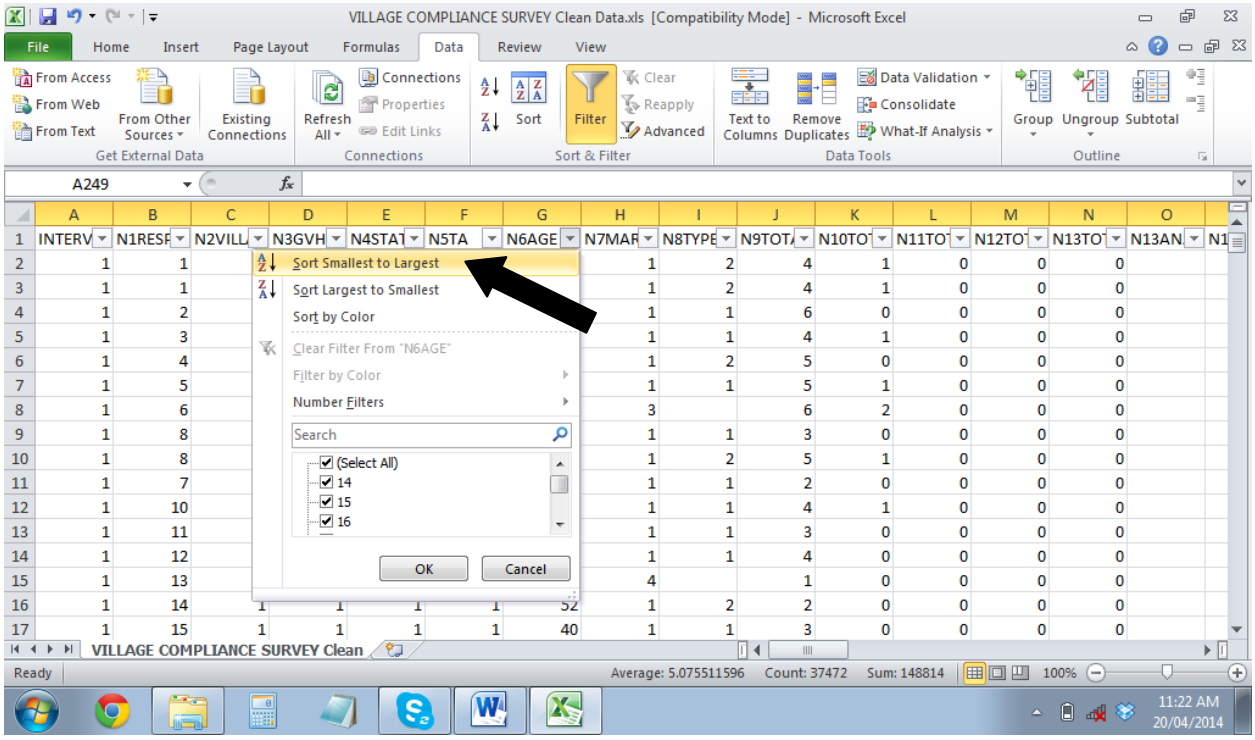
Step 3: Click on the Data menu item. Then click the Filter button.

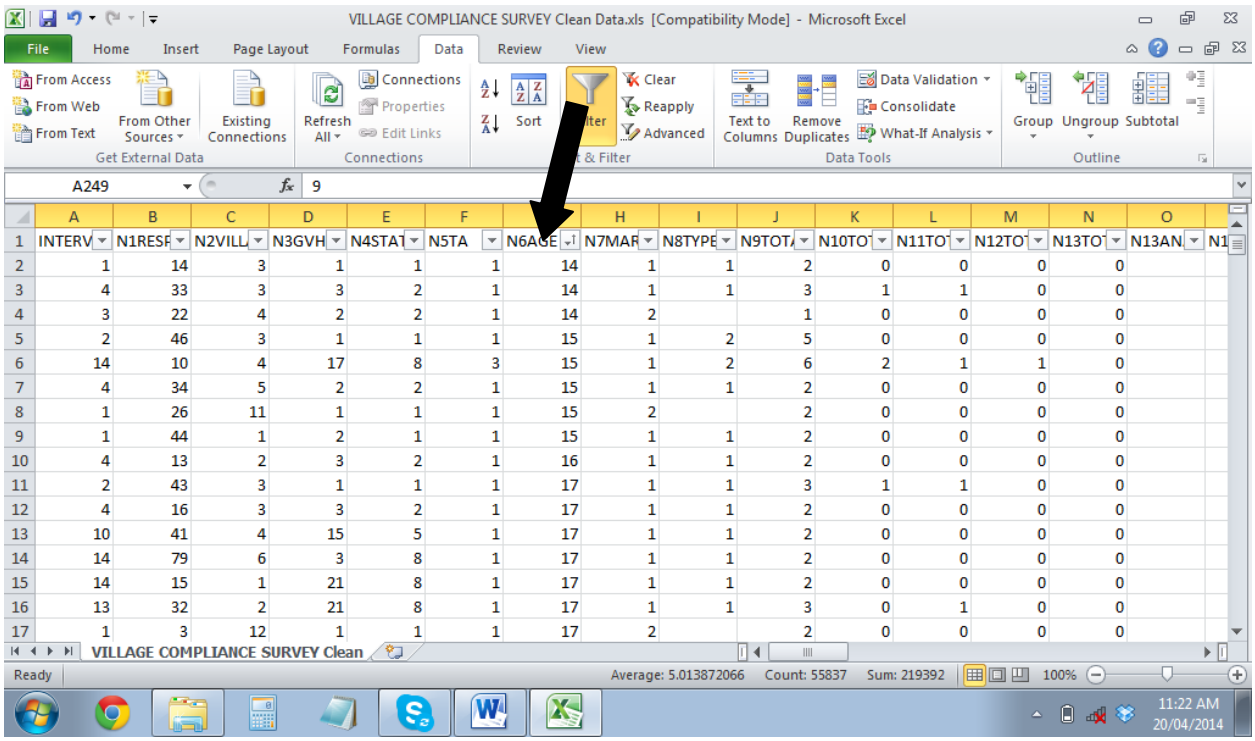


Step 4: Small arrows should appear at the top of each column.

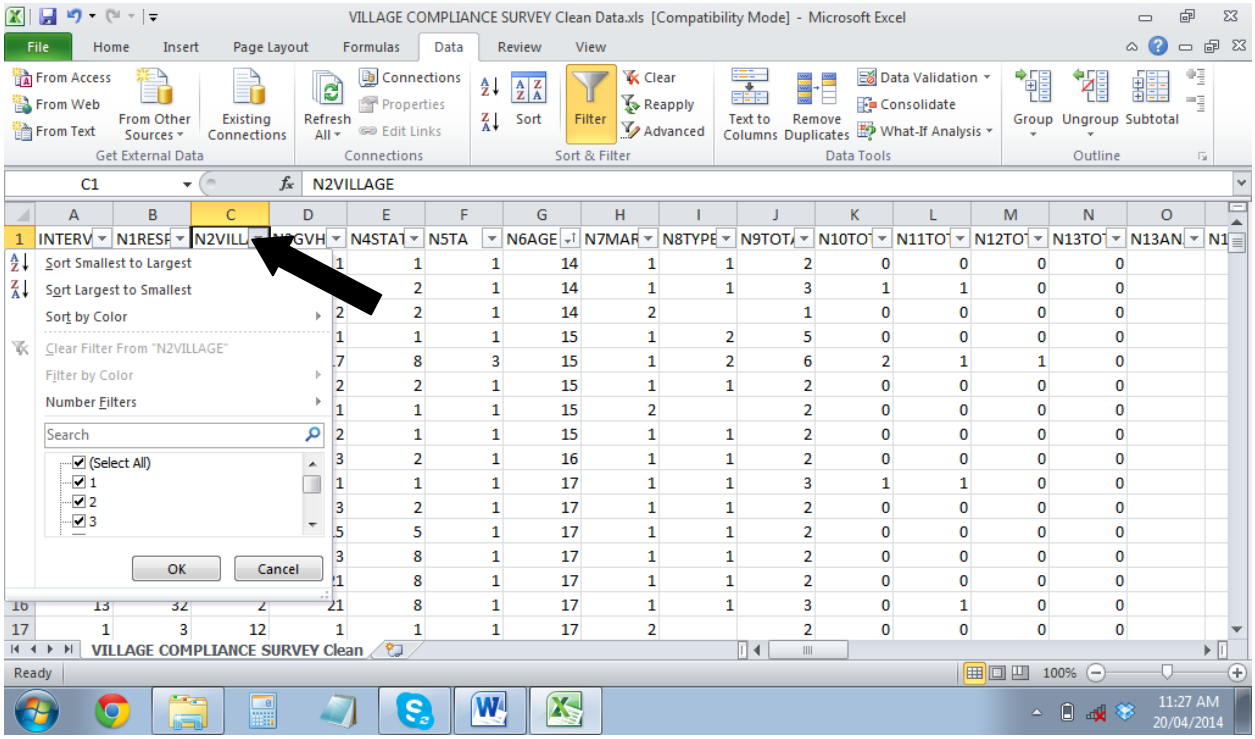


Step 5: If you click on one of the filter arrows it will show you everything that has been entered into that column. You can choose to sort the column from smallest to largest or largest to smallest. In the example below I have sorted the rows based on the Age column from youngest to oldest.

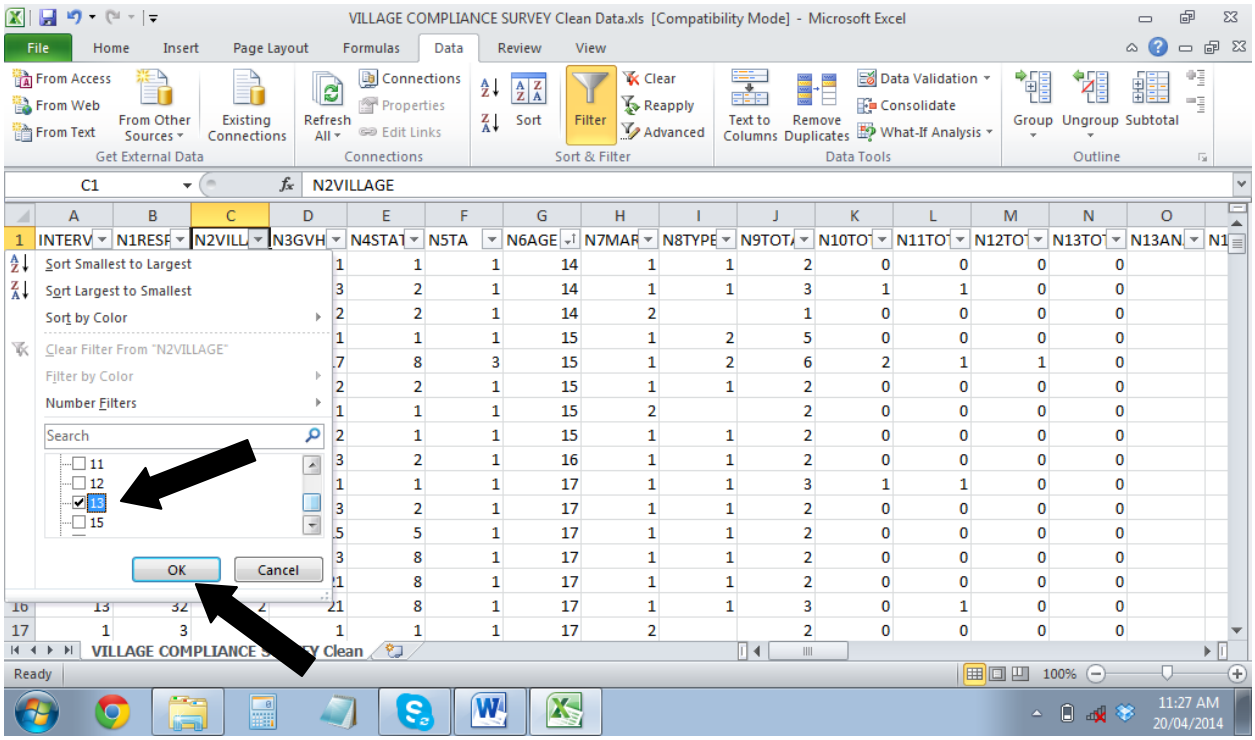




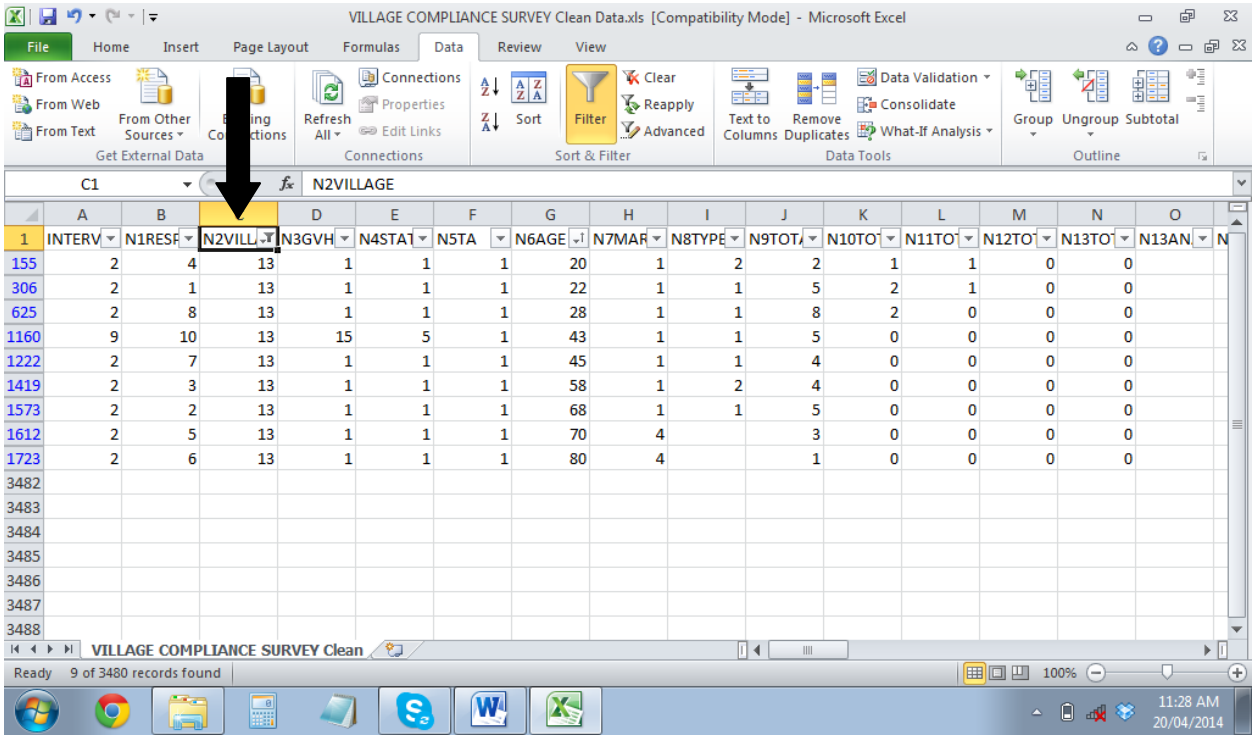
Step 6: You can also choose to see only some of the rows. For example, to see only the rows for a particular GVH or village. To do this click on the filter button for the column that you want to use (e.g. village code).



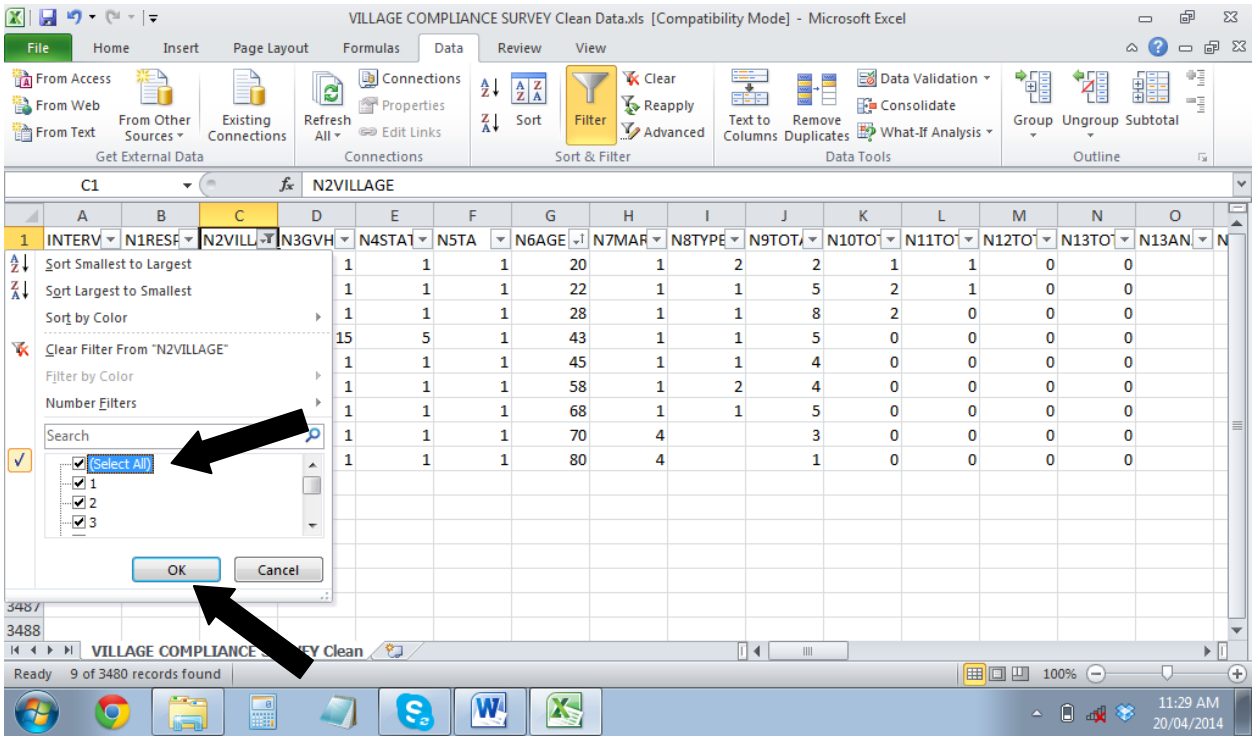
Step 7: Then click Select All to remove all the ticks. Tick only the items that you want to see (in this case I want to see all households from village #13) and click ok.

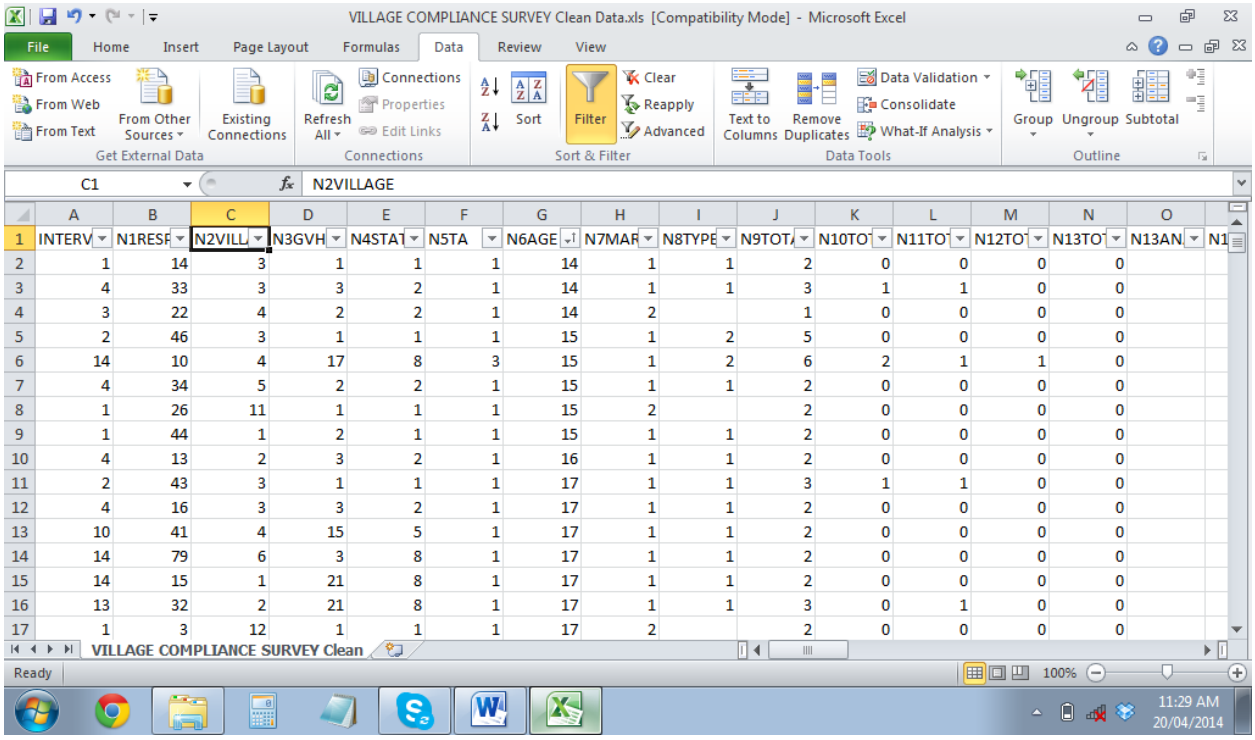


Step 8: Only the rows that you have selected will appear (in this case rows for village #13).



Step 9: To see all the rows again you need to remove the filter. First click on the filter button. Then tick Select All and click OK.





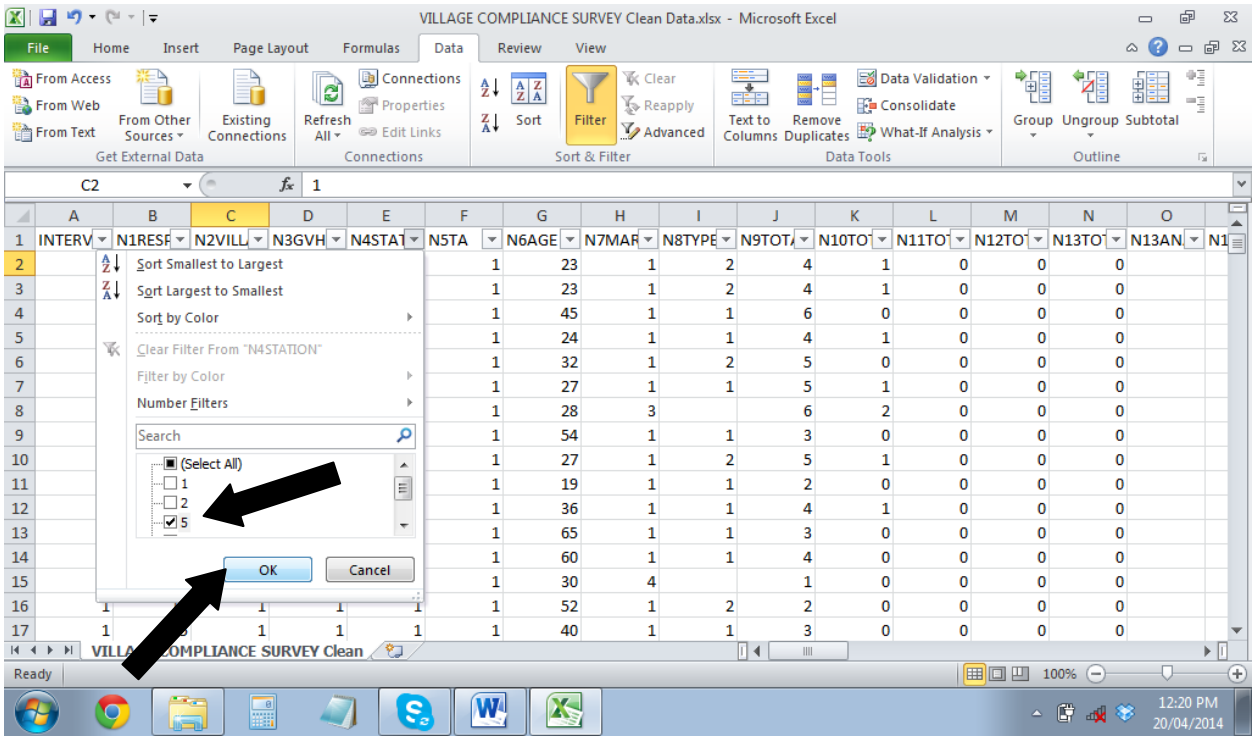
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	INTERV	N1RESP	N2VILL	N3GVH	N4STAT	N5TA	N6AGE	N7MAR	N8TYPE	N9TOT	N10TO	N11TO	N12TO	N13TO	N13AN
2	1	14	3	1	1	1	14	1	1	2	0	0	0	0	
3	4	33	3	3	2	1	14	1	1	3	1	1	0	0	
4	3	22	4	2	2	1	14	2		1	0	0	0	0	
5	2	46	3	1	1	1	15	1	2	5	0	0	0	0	
6	14	10	4	17	8	3	15	1	2	6	2	1	1	0	
7	4	34	5	2	2	1	15	1	1	2	0	0	0	0	
8	1	26	11	1	1	1	15	2		2	0	0	0	0	
9	1	44	1	2	1	1	15	1	1	2	0	0	0	0	
10	4	13	2	3	2	1	16	1	1	2	0	0	0	0	
11	2	43	3	1	1	1	17	1	1	3	1	1	0	0	
12	4	16	3	3	2	1	17	1	1	2	0	0	0	0	
13	10	41	4	15	5	1	17	1	1	2	0	0	0	0	
14	14	79	6	3	8	1	17	1	1	2	0	0	0	0	
15	14	15	1	21	8	1	17	1	1	2	0	0	0	0	
16	13	32	2	21	8	1	17	1	1	3	0	1	0	0	
17	1	3	12	1	1	1	17	2		2	0	0	0	0	

4.2 Auditing data entry

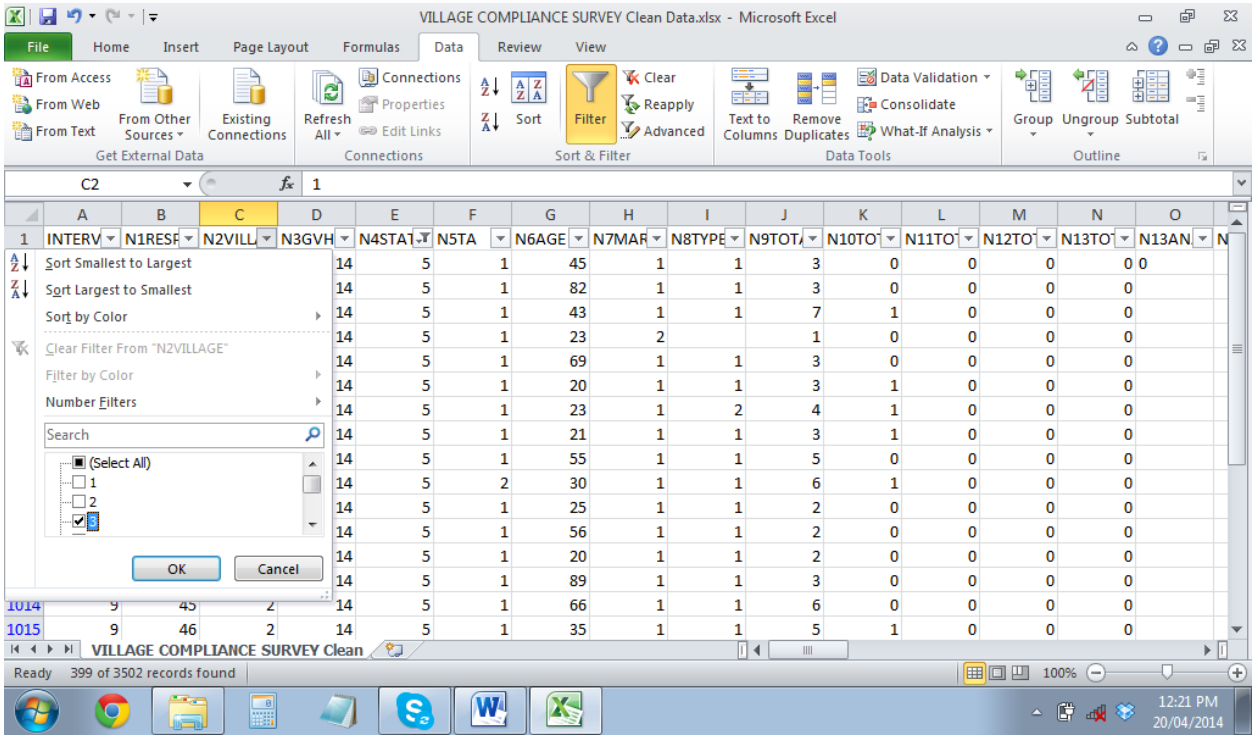
After setting up the filter you need to audit the data entry to make sure it was done accurately.

Step 1: Randomly choose some hard copy surveys. The number you choose depends on how many surveys you have in total, and how thoroughly you want to audit. A good starting point is to audit 5-10% of all the surveys.

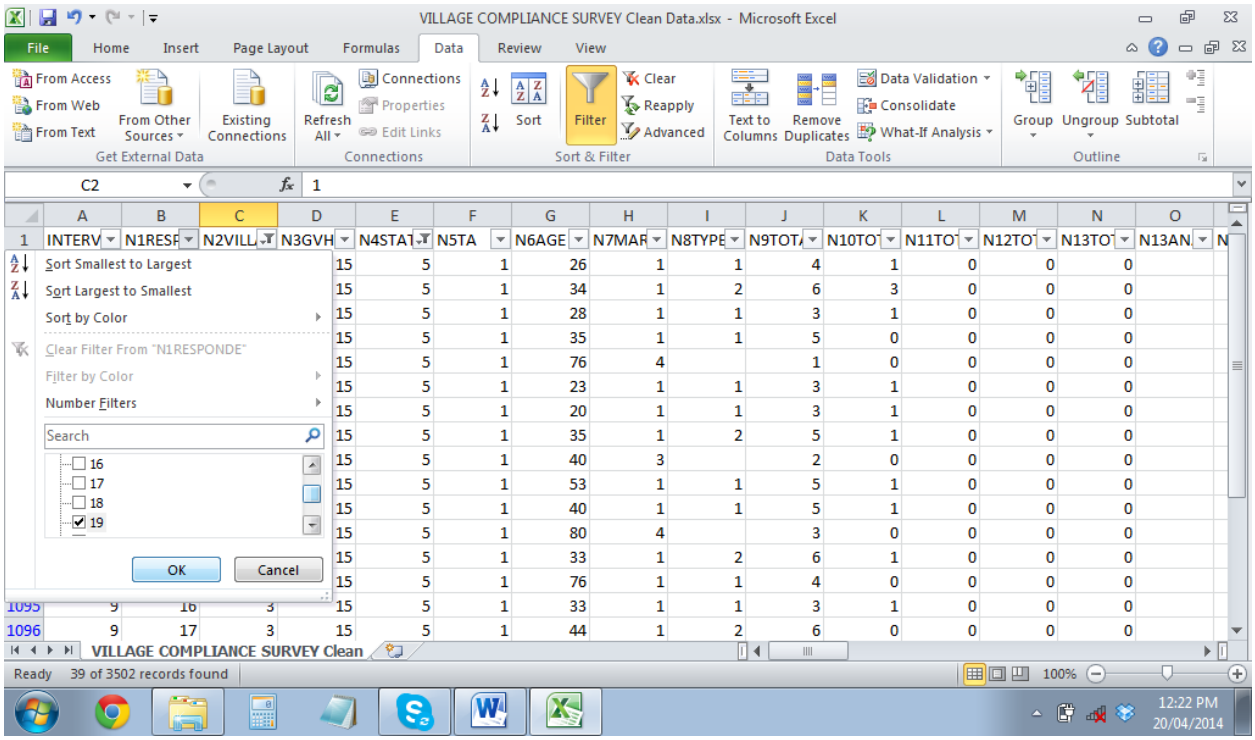
Step 2: For each survey you need to find the corresponding row in the spreadsheet using filters. First filter by the station. For example, if the survey is from Nyamazani (code #5) then use the filter to show only those surveys from Nyamazani. Click on the filter button at the top of the station column. Click on Select All to remove the ticks, then tick the code for the station (in this case 5). Click OK.



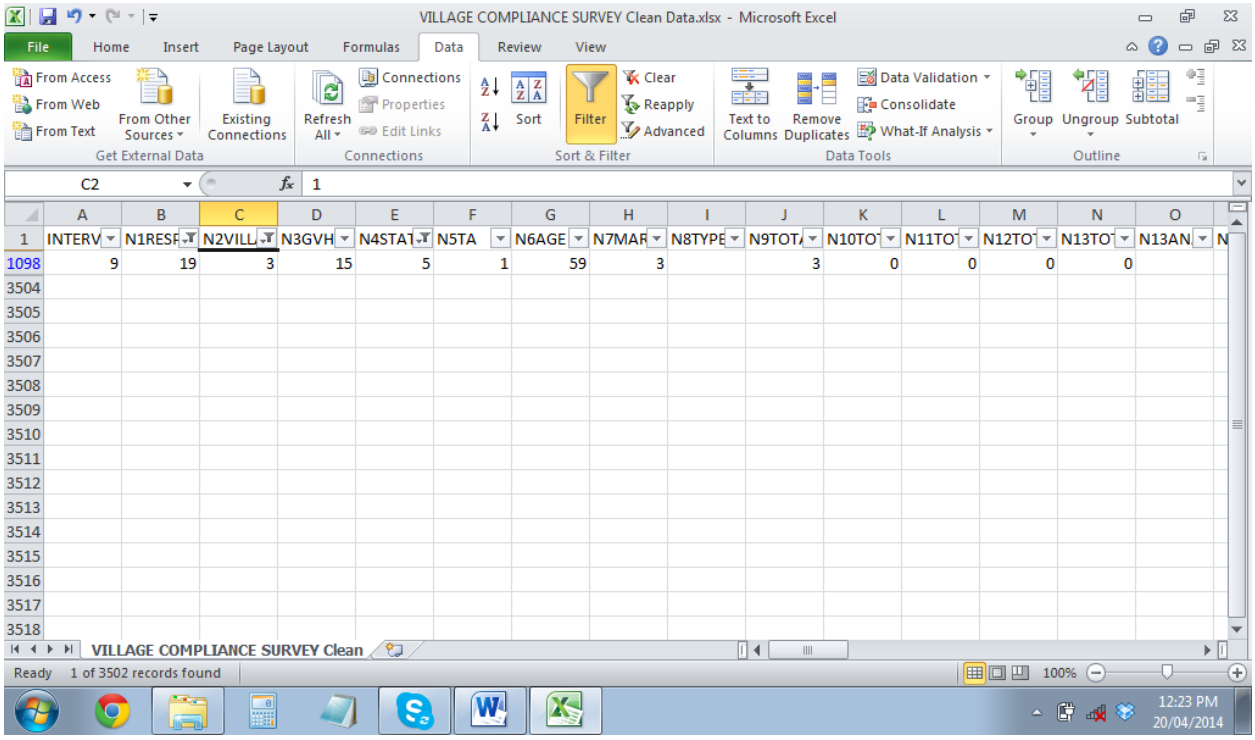
Step 3: Only the rows for the station you have selected will appear. Then use the same steps to filter by the village. For example, Gwinze, which has the code #3.



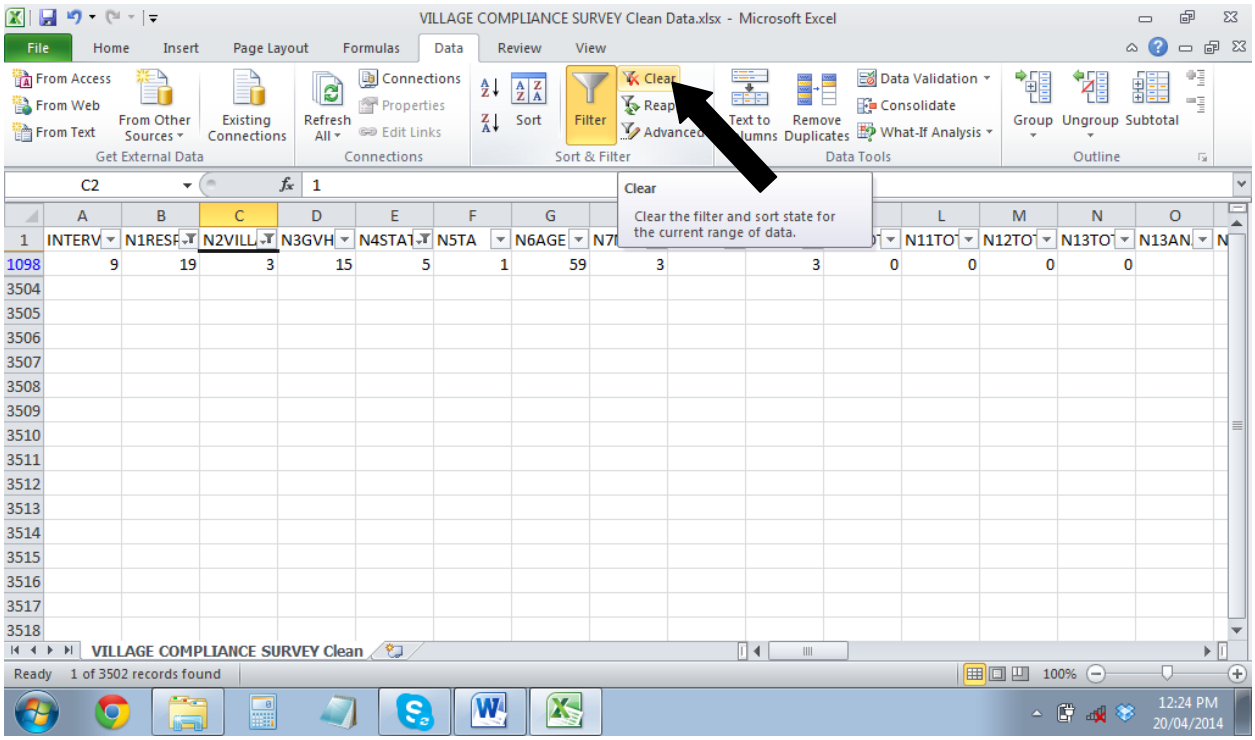
Step 4: Only the rows for the village you have selected will appear. Use the same steps again to filter by respondent. For example, respondent #19.

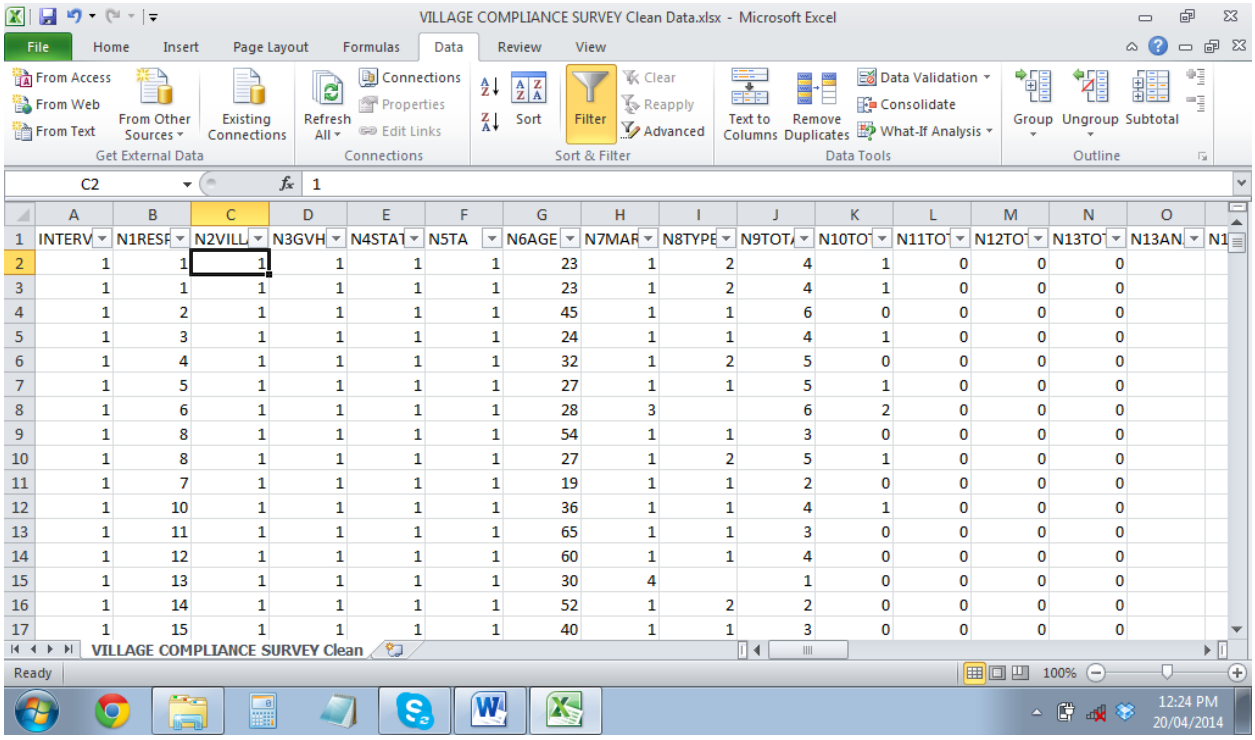


Step 5: Now there should be only one row appearing (in this case respondent #13 from village #3 in station #5). Choose 10 questions in the hard copy survey and compare them to the data in the spreadsheet. Make a note of any mistakes in the data and correctly them immediately on the spreadsheet.



Step 6: When you have finished auditing the survey click the Clear button on the top menu bar to clear all the filters. All the rows will appear again. Start again at Step 2 for the next survey.





	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	INTERV	N1RESP	N2VILL	N3GVH	N4STAT	N5TA	N6AGE	N7MAR	N8TYPE	N9TOT	N10TO	N11TO	N12TO	N13TO	N13AN
2	1	1	1	1	1	1	23	1	2	4	1	0	0	0	
3	1	1	1	1	1	1	23	1	2	4	1	0	0	0	
4	1	2	1	1	1	1	45	1	1	6	0	0	0	0	
5	1	3	1	1	1	1	24	1	1	4	1	0	0	0	
6	1	4	1	1	1	1	32	1	2	5	0	0	0	0	
7	1	5	1	1	1	1	27	1	1	5	1	0	0	0	
8	1	6	1	1	1	1	28	3		6	2	0	0	0	
9	1	8	1	1	1	1	54	1	1	3	0	0	0	0	
10	1	8	1	1	1	1	27	1	2	5	1	0	0	0	
11	1	7	1	1	1	1	19	1	1	2	0	0	0	0	
12	1	10	1	1	1	1	36	1	1	4	1	0	0	0	
13	1	11	1	1	1	1	65	1	1	3	0	0	0	0	
14	1	12	1	1	1	1	60	1	1	4	0	0	0	0	
15	1	13	1	1	1	1	30	4		1	0	0	0	0	
16	1	14	1	1	1	1	52	1	2	2	0	0	0	0	
17	1	15	1	1	1	1	40	1	1	3	0	0	0	0	

Step 7: When you have finished auditing all the surveys count the total number of questions you audited (10 questions multiplied by the number of surveys audited). Also count the total number of questions that had an error.

Divide the number of correct questions by the total number of questions audited and multiply by 100 to get the accuracy rate for data entry. If the accuracy rate is below 90% consider checking all the surveys again to correct errors.

Example:

I audited 10 questions in 15 surveys.

The total number of questions I audited was 10 x 15 = 150 questions.

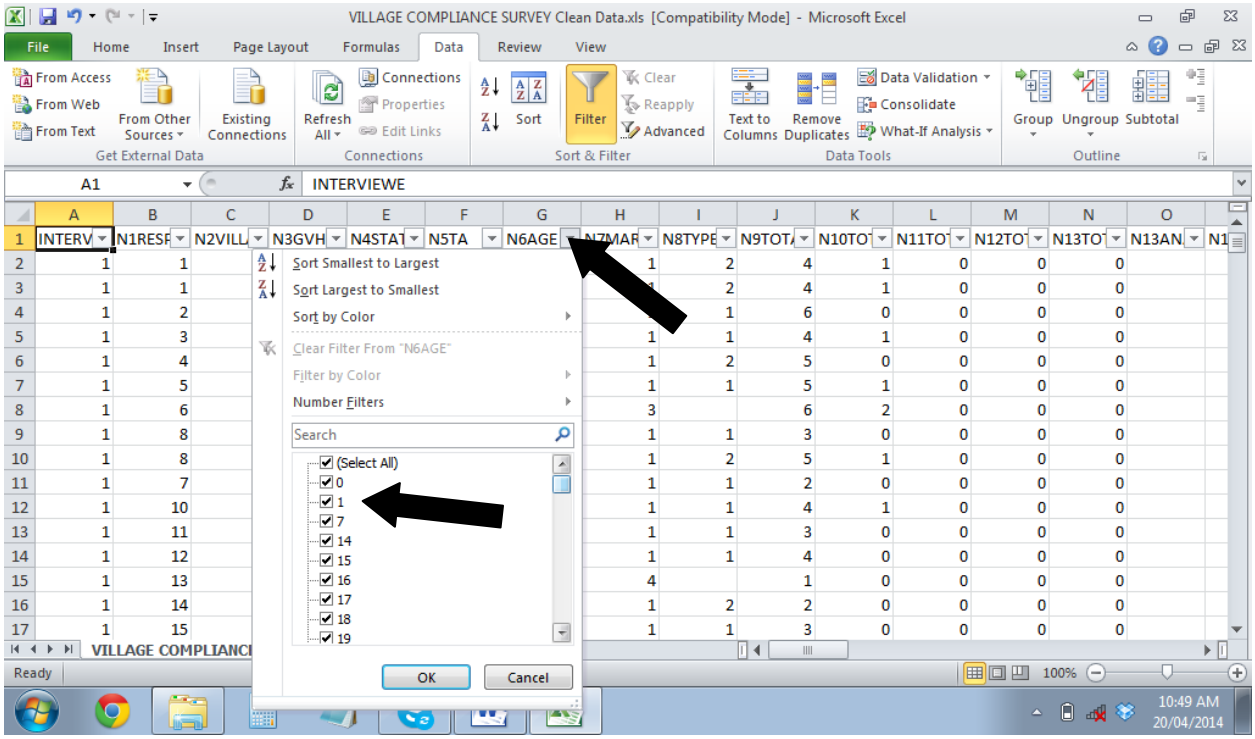
Out of 150 questions 23 had an error. That means 150 – 23 = 127 were correct.

Therefore, the accuracy rate for data entry was 127 / 150 x 100 = 84%

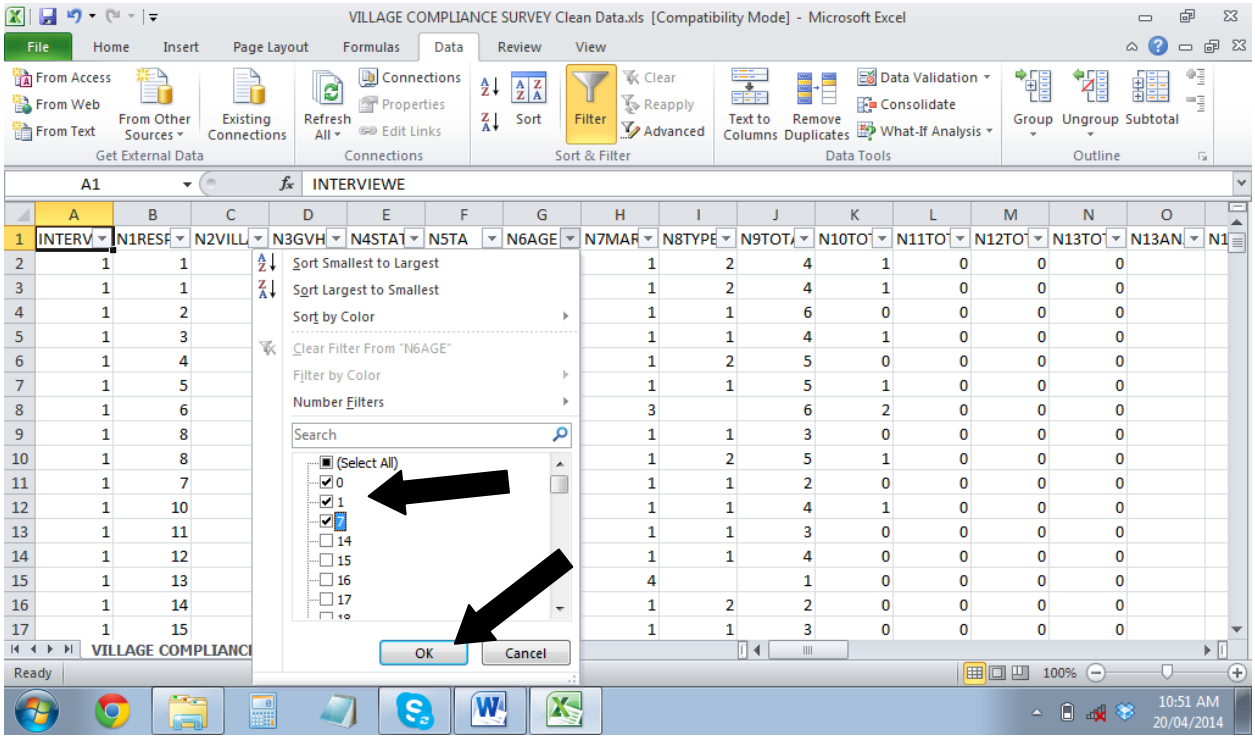
4.3 Checking each column

After auditing the data entry need to use the filter options to check each column in the spreadsheet for unusual data. If you find anything unusual you should correct it before proceeding to the analysis.

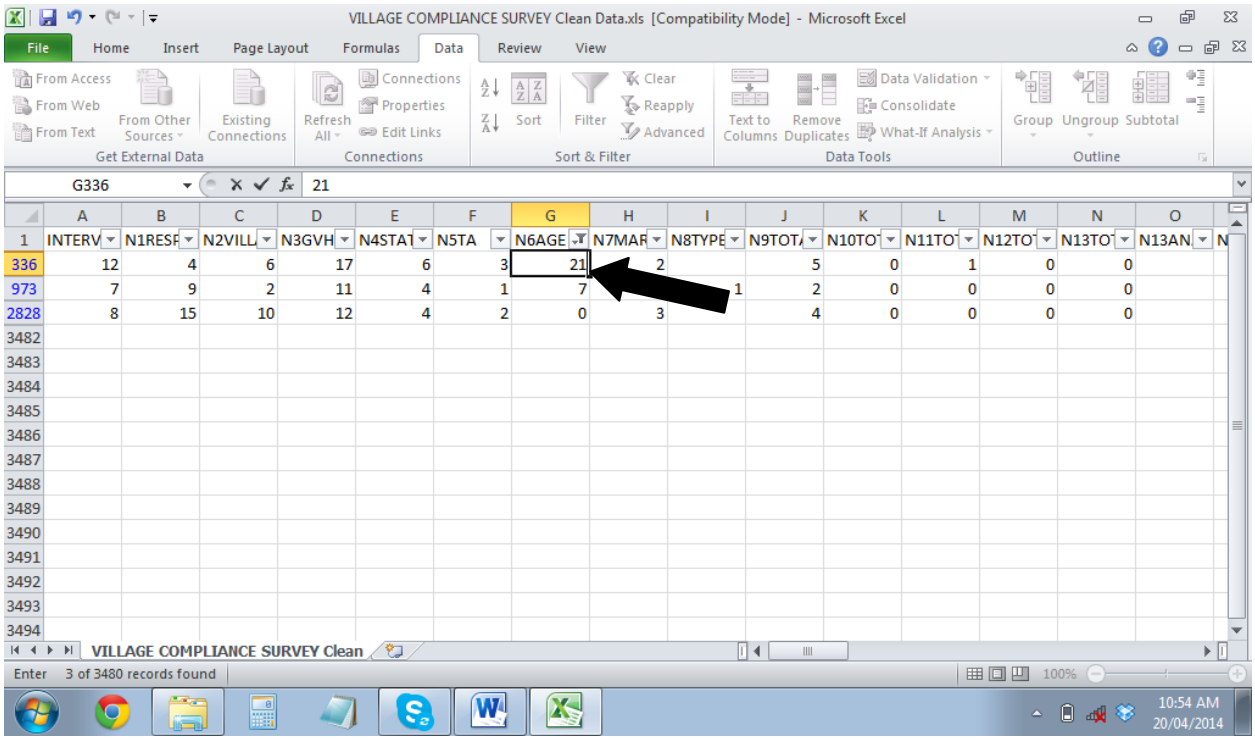
Step 1: Click on one of the filter arrows to show you everything that has been entered into that column. Check the list of results to make sure there is nothing usual. For example, if I click on the Age column it shows me that for some surveys the age was entered as 0, 1 and 7. This is not possible, so it needs to be corrected.



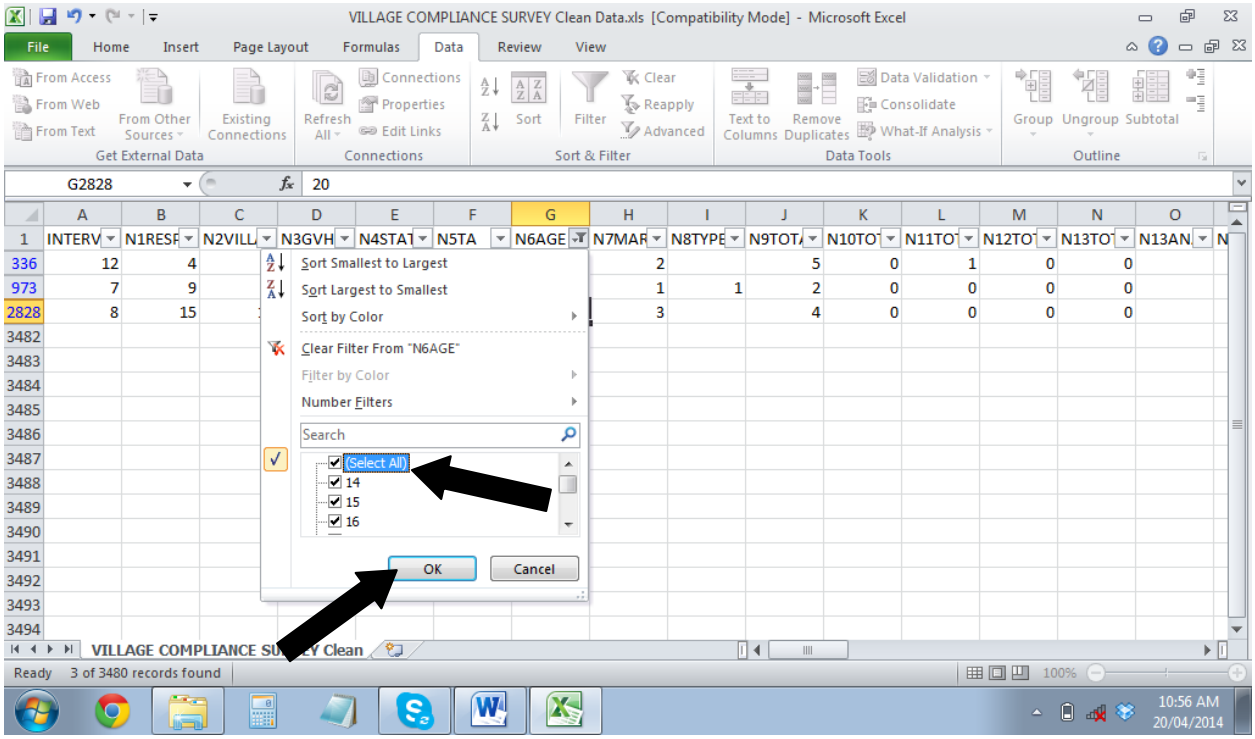
Step 2: To correct the error click Select All to remove all the ticks. Then tick only the items that are a problem (in this case 0, 1 and 7) and click OK.



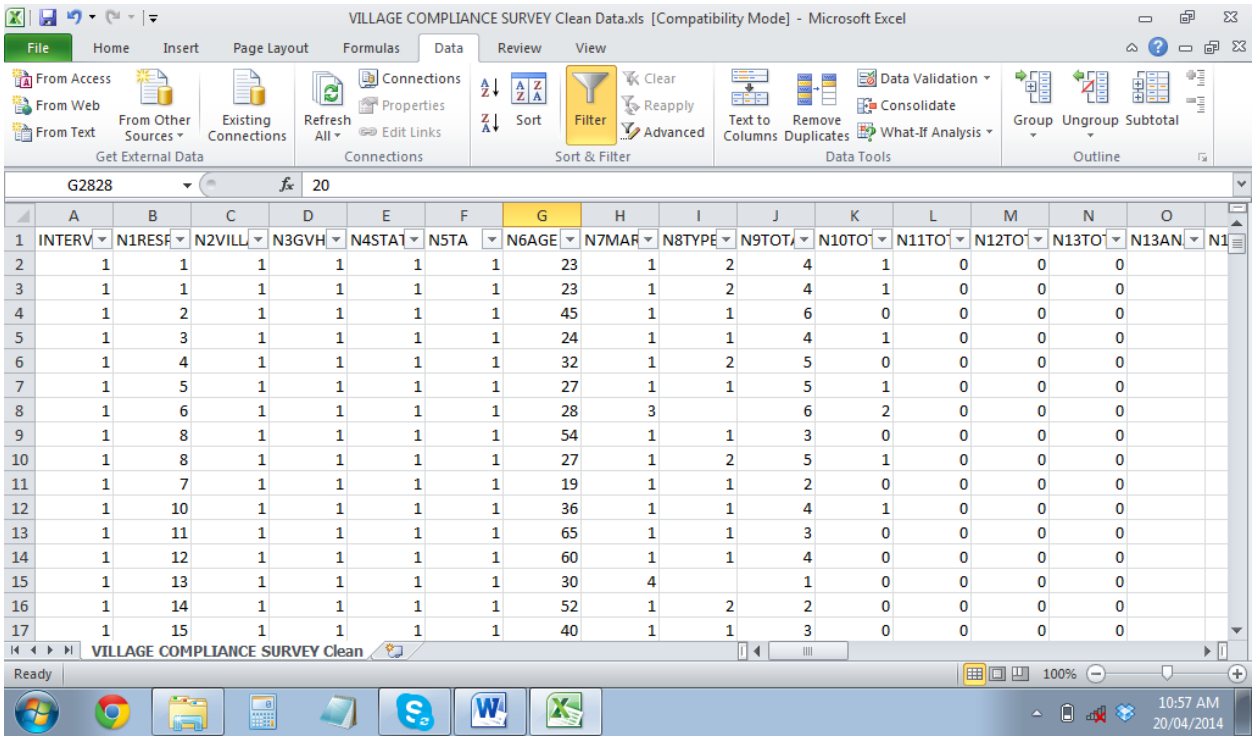
Step 7: The result is that only the rows with the errors appear. You can now correct the errors by finding the original hard copy of the survey and typing the correct number directly into the cell.



Step 7: Once you have finished correcting all the errors in a column click on the filter arrow again. All the problem numbers should have disappeared. Click Select All to tick all the boxes then click OK.



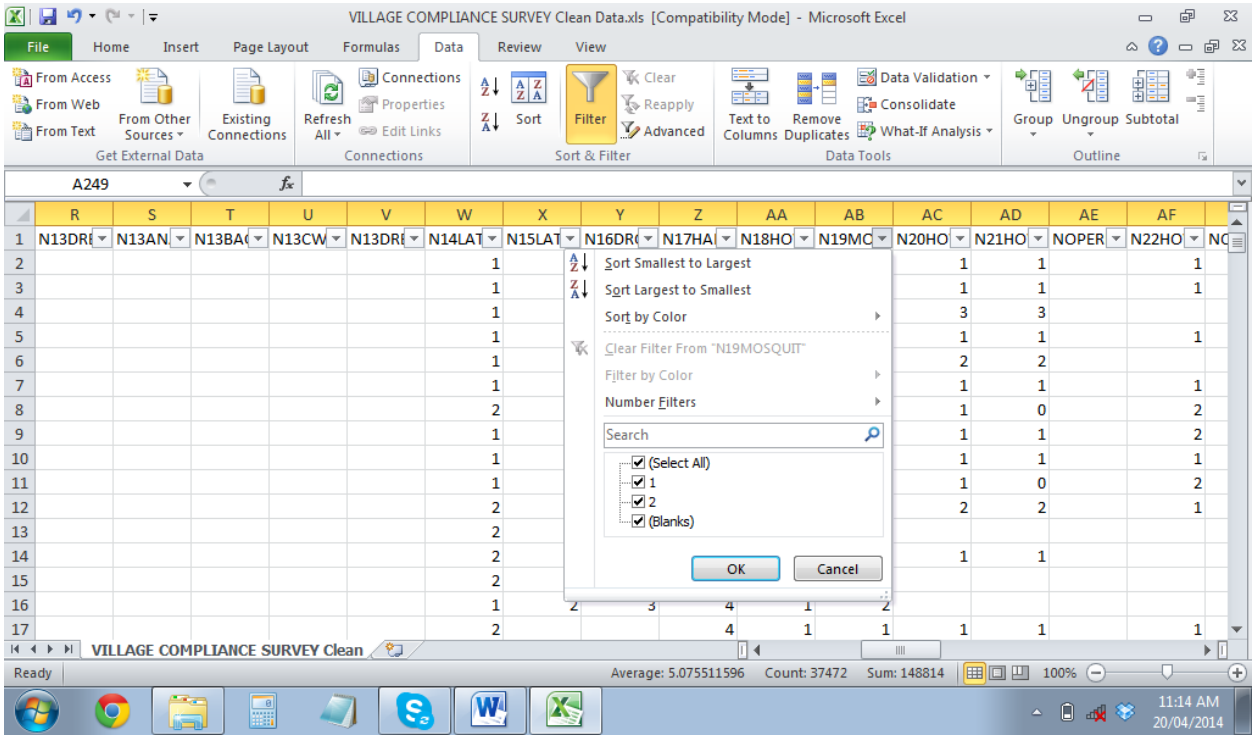
Step 8: All the surveys should appear again. You can now move on to checking the next column.



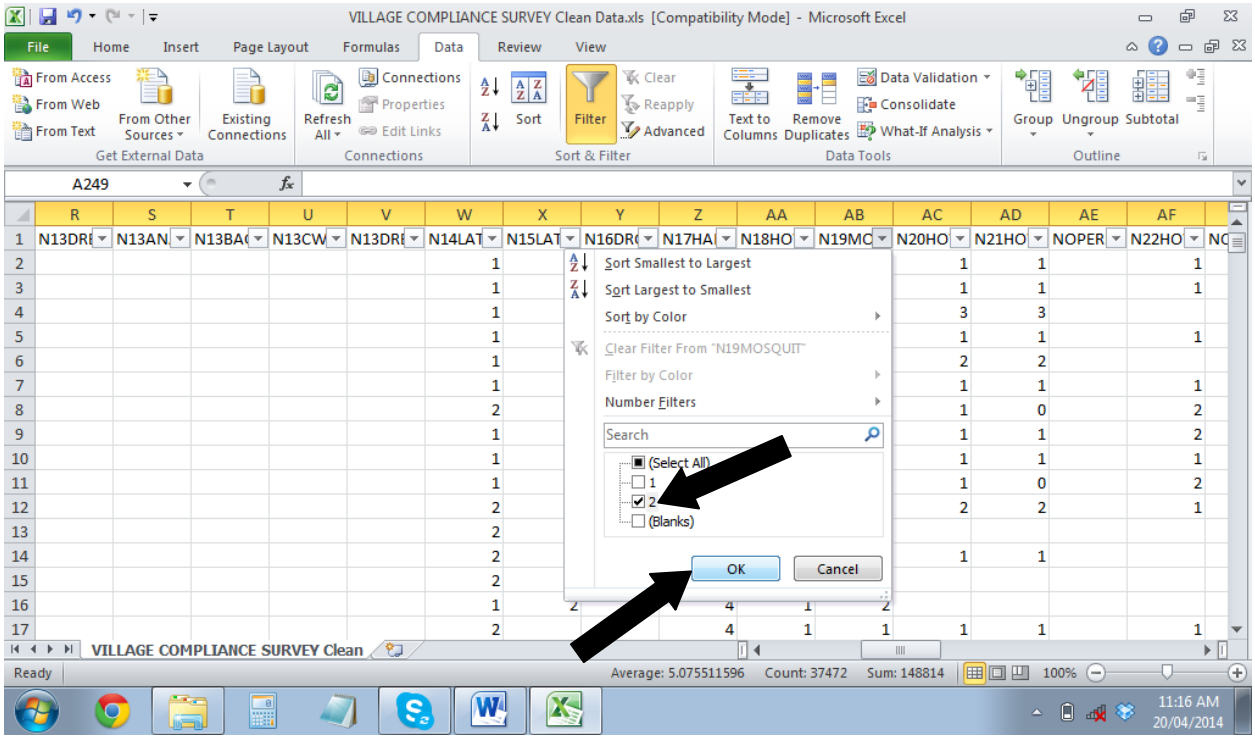
4.4 Checking jumps

If the survey has any jumps between questions you need to check that the jump has also been made in the data entry.

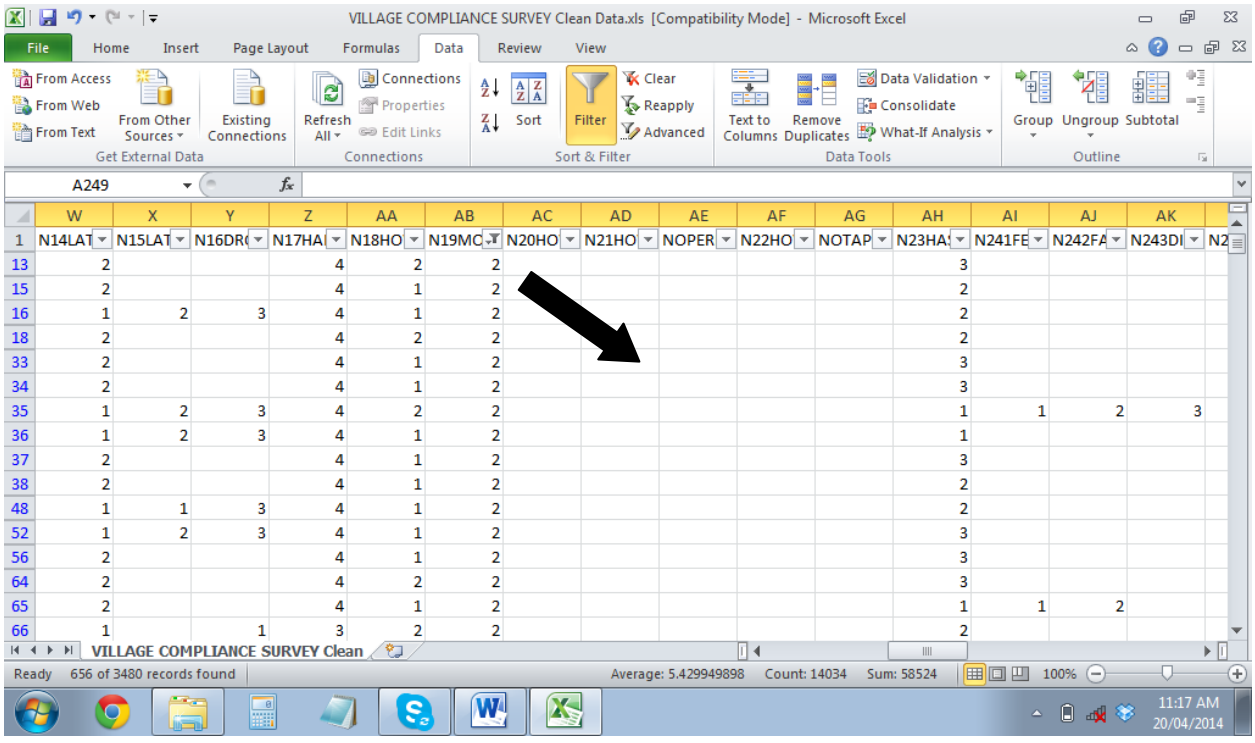
Step 1: Find the column that corresponds to the question with a jump. For example, if *Q19 Does the household have a mosquito net?* jumps to Q23 if the answer is *No* then go to Q19. Click on the filter button for that column.



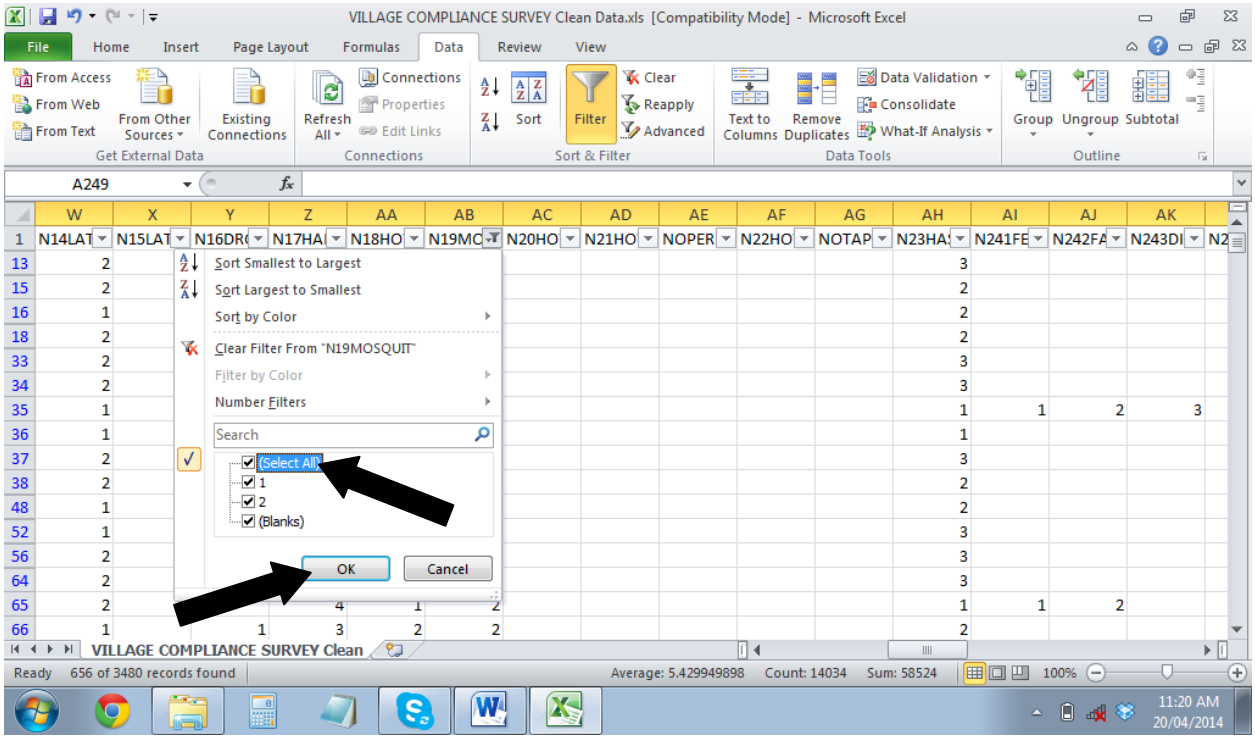
Step 2: Click Select All to remove all the tick. Then tick on the answer that would result in a jump (in this case answer is No = 2) and click OK.



Step 3: Now you will only see rows where the answer to Q19 was 2 (No). Check the columns that should have been jumped (in this case Q20, Q21 and Q22) to make sure they are all blank. If any of them are not blank then find the original hard copy of the survey and make a correction.



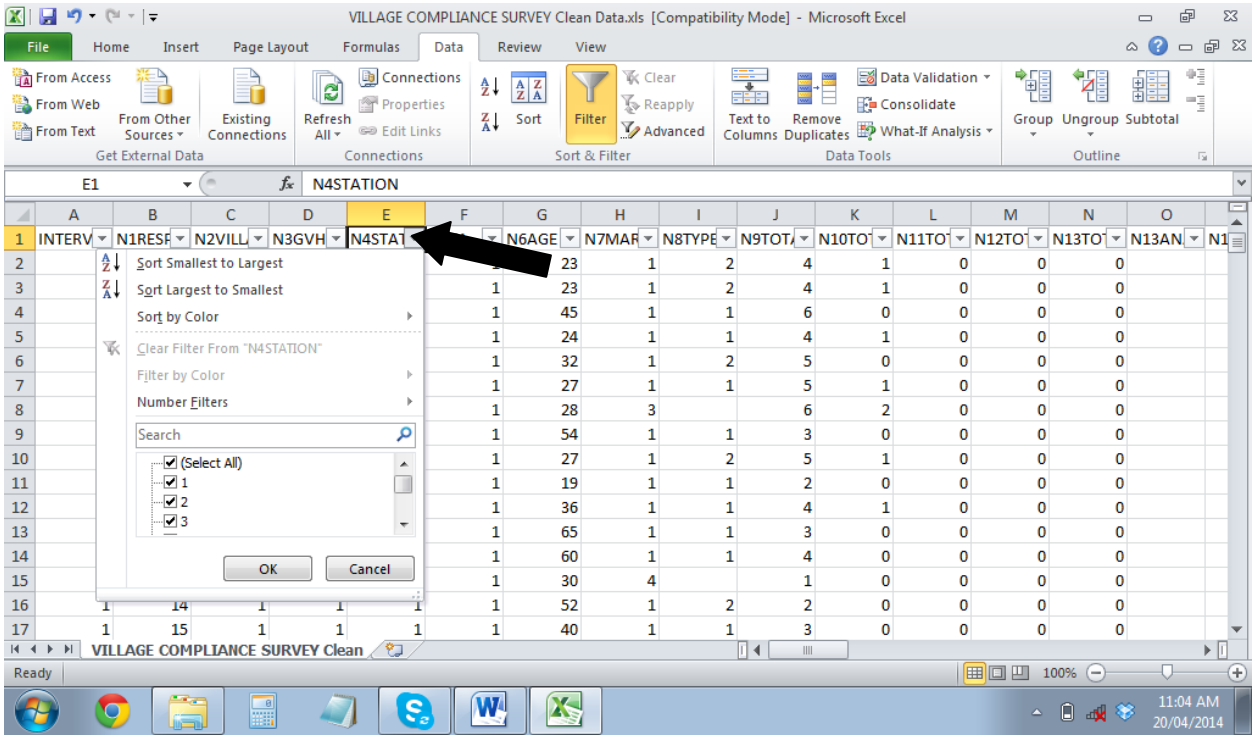
Step 4: Once all the corrections have been made click on the filter button again. Click Select All to tick all the items then click OK. All the rows will now return. Move on to the next question with a jump.


VILLAGE COMPLIANCE SURVEY Clean Data.xls [Compatibility Mode] - Microsoft Excel
File
Home
Insert
Page Layout
Formulas
Data
Review
View
From Access
From Web
From Text
From Other Sources
Existing Connections
Get External Data
Refresh All
Connections
Properties
Edit Links
Connections
Sort
Filter
Clear
Reapply
Advanced
Sort & Filter
Text to Columns
Remove Duplicates
Data Validation
Consolidate
What-If Analysis
Data Tools
Group
Ungroup
Subtotal
Outline
A249
Sort Smallest to Largest
Sort Largest to Smallest
Sort by Color
Clear Filter From "N19MOSQUIT"
Filter by Color
Number Filters
Search
(Select All)
1
2
(Blanks)
OK
Cancel
VILLAGE COMPLIANCE SURVEY Clean
Ready
656 of 3480 records found
Average: 5.429949898
Count: 14034
Sum: 58524
100%
11:20 AM
20/04/2014

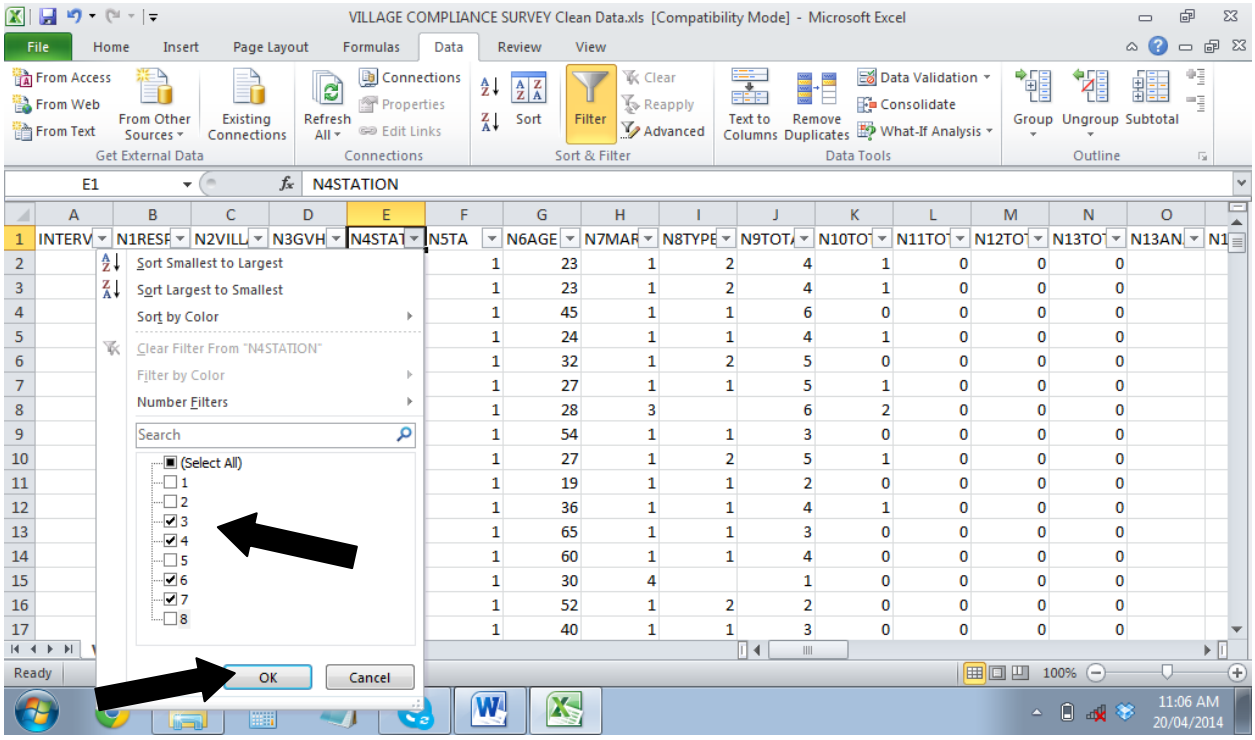
4.5 Deleting rows

Sometimes you may want to delete particular rows from the data set. For example, if you only want to do analysis for one station you need to delete the rows for the other stations.

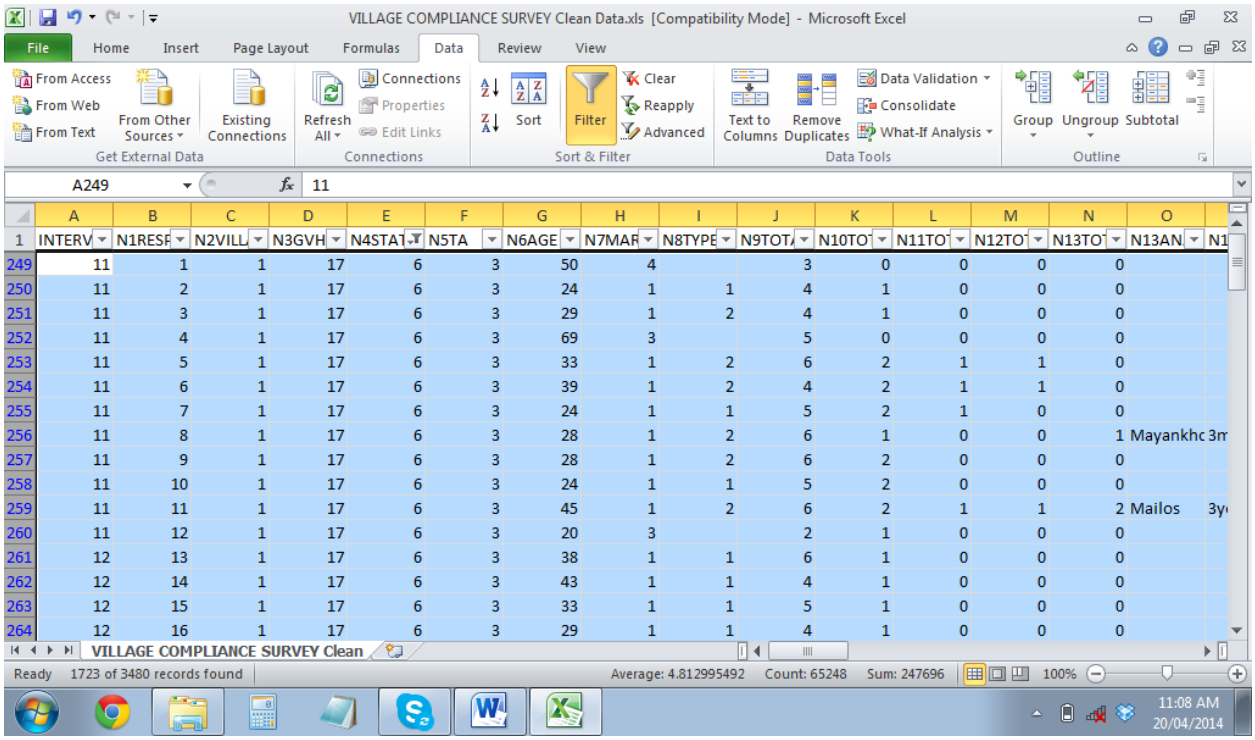
Step 1: Go to the column that you want to use to select rows for deleting. For example, if you want to delete rows based on the station then do to the station code column. If you want to delete rows based on the village then go to the village code column. In this example we will use the station. Click on the filter button for that column.

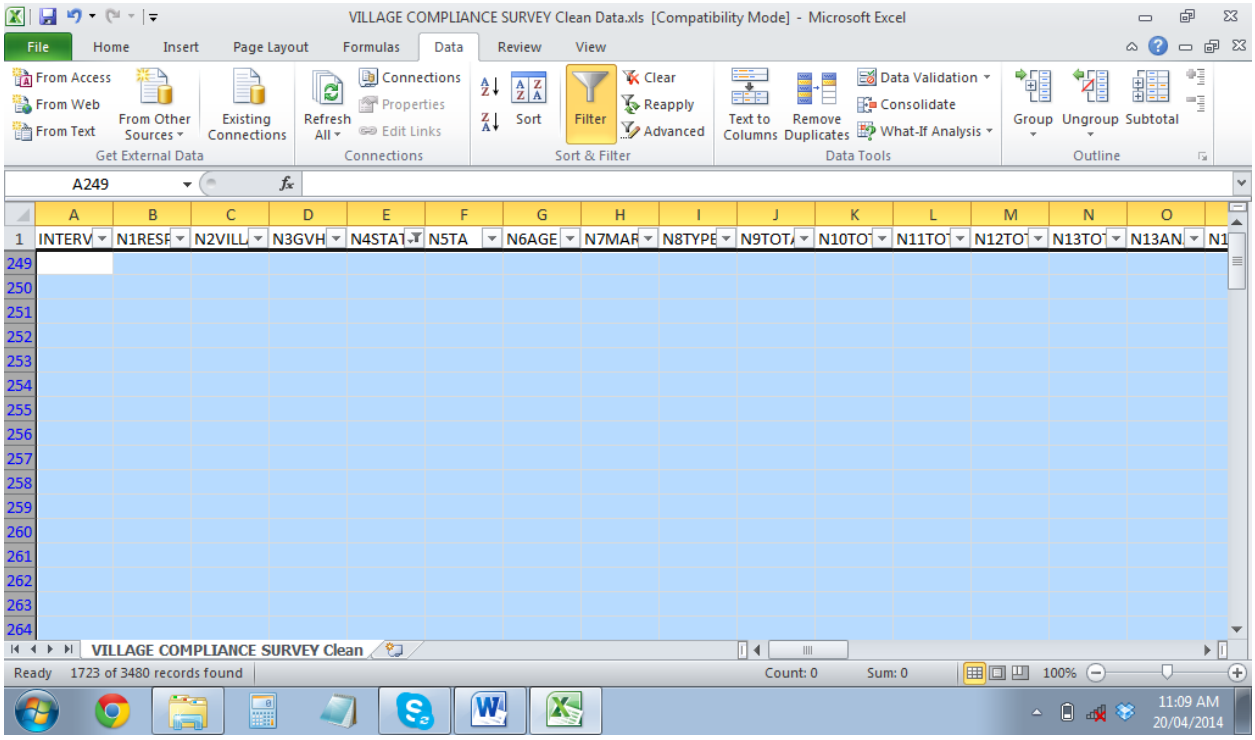


Step 2: Click Select All to remove all the ticks. Then tick only the ones that you want to delete. In this case I want to delete all surveys from station codes 3, 4, 6 and 7, so I tick those numbers and click OK.

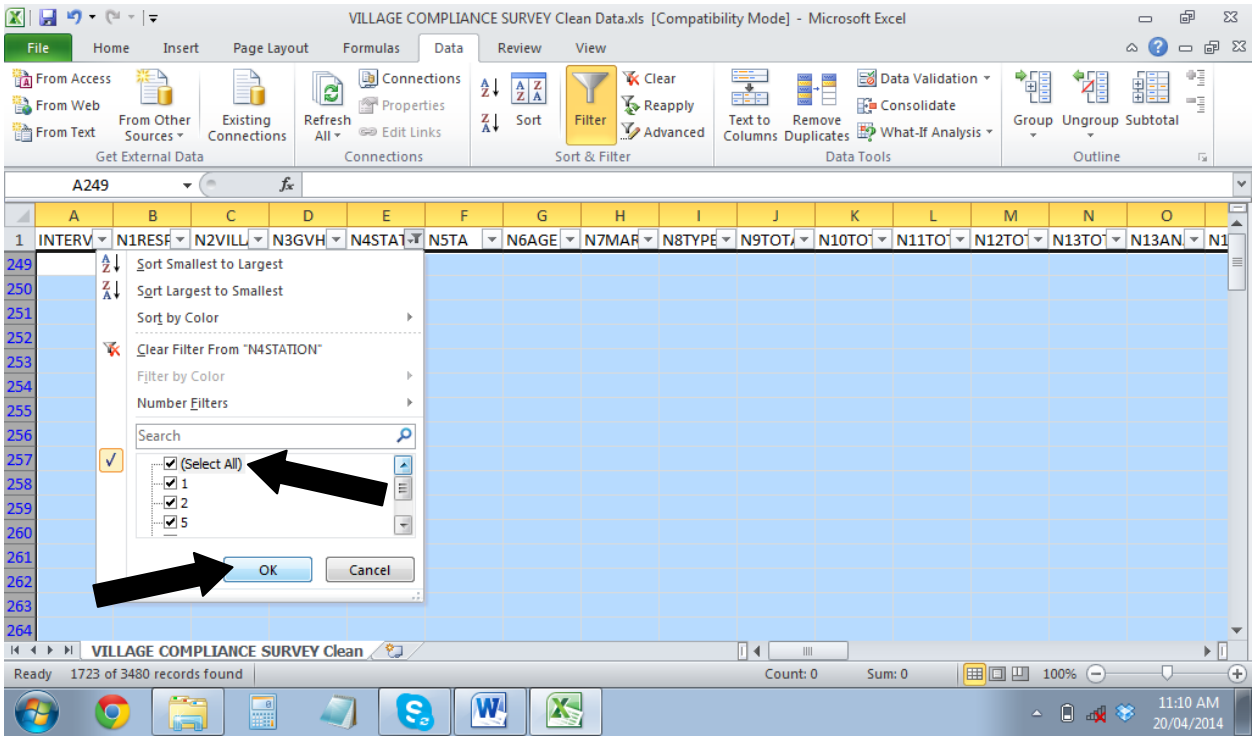


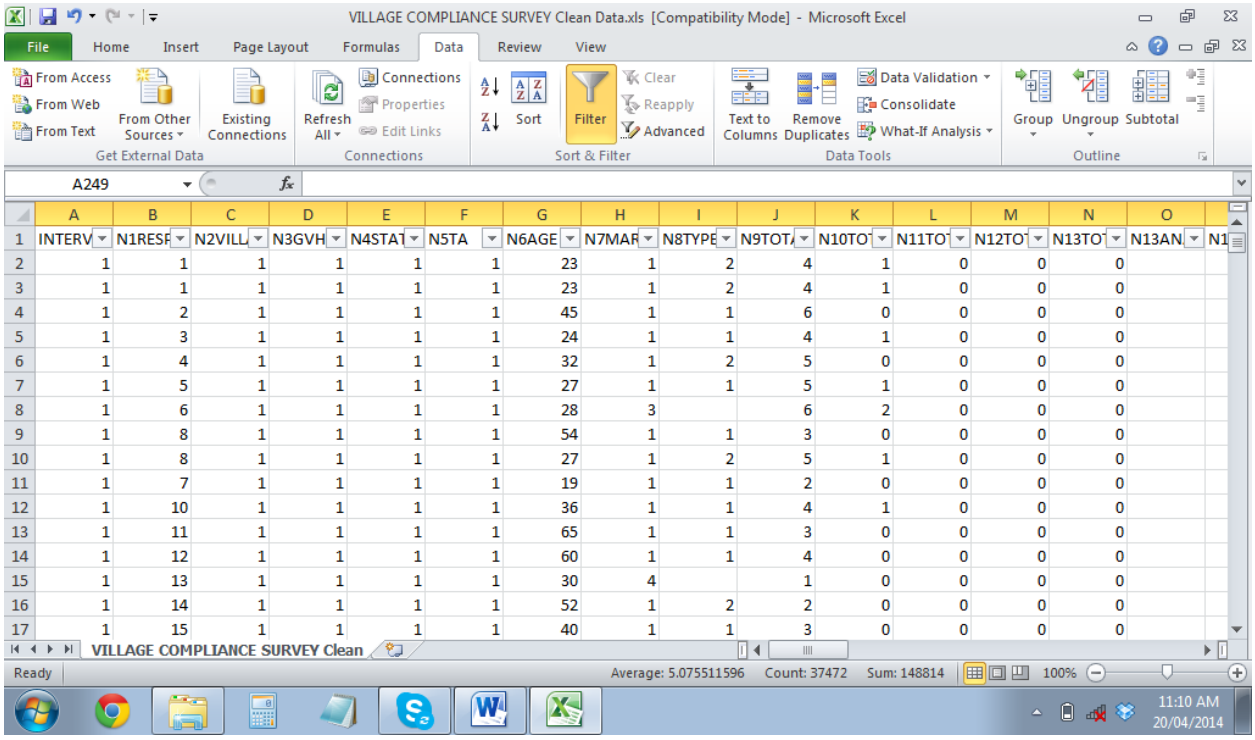
Step 3: Only the rows that you want to delete will appear. Select all the rows then press the Delete key on the keyboard.





Step 4: Click on the filter button again. The code numbers should have disappeared. Choose Select All and click OK. The remaining rows should now appear.





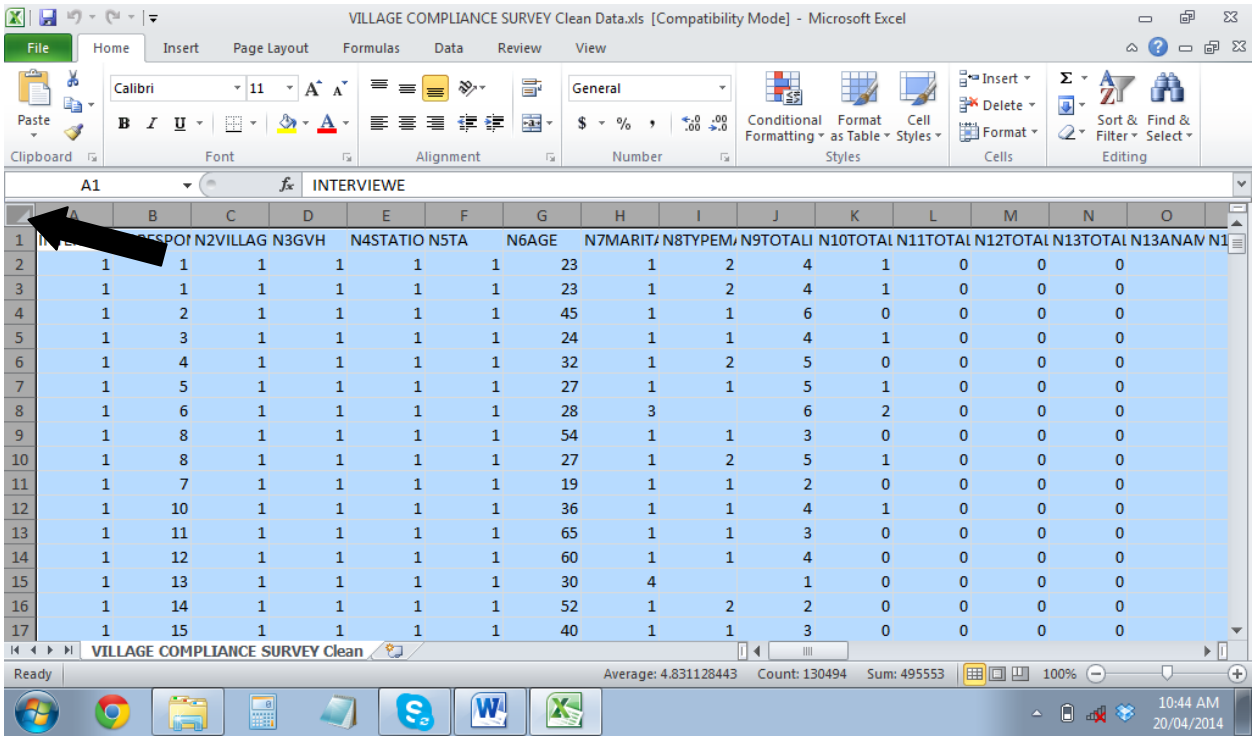
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	INTERV	N1RESP	N2VILL	N3GVH	N4STAT	N5TA	N6AGE	N7MAR	N8TYPE	N9TOT	N10TO	N11TO	N12TO	N13TO	N13AN
2	1	1	1	1	1	1	23	1	2	4	1	0	0	0	
3	1	1	1	1	1	1	23	1	2	4	1	0	0	0	
4	1	2	1	1	1	1	45	1	1	6	0	0	0	0	
5	1	3	1	1	1	1	24	1	1	4	1	0	0	0	
6	1	4	1	1	1	1	32	1	2	5	0	0	0	0	
7	1	5	1	1	1	1	27	1	1	5	1	0	0	0	
8	1	6	1	1	1	1	28	3		6	2	0	0	0	
9	1	8	1	1	1	1	54	1	1	3	0	0	0	0	
10	1	8	1	1	1	1	27	1	2	5	1	0	0	0	
11	1	7	1	1	1	1	19	1	1	2	0	0	0	0	
12	1	10	1	1	1	1	36	1	1	4	1	0	0	0	
13	1	11	1	1	1	1	65	1	1	3	0	0	0	0	
14	1	12	1	1	1	1	60	1	1	4	0	0	0	0	
15	1	13	1	1	1	1	30	4		1	0	0	0	0	
16	1	14	1	1	1	1	52	1	2	2	0	0	0	0	
17	1	15	1	1	1	1	40	1	1	3	0	0	0	0	

5 Data Analysis

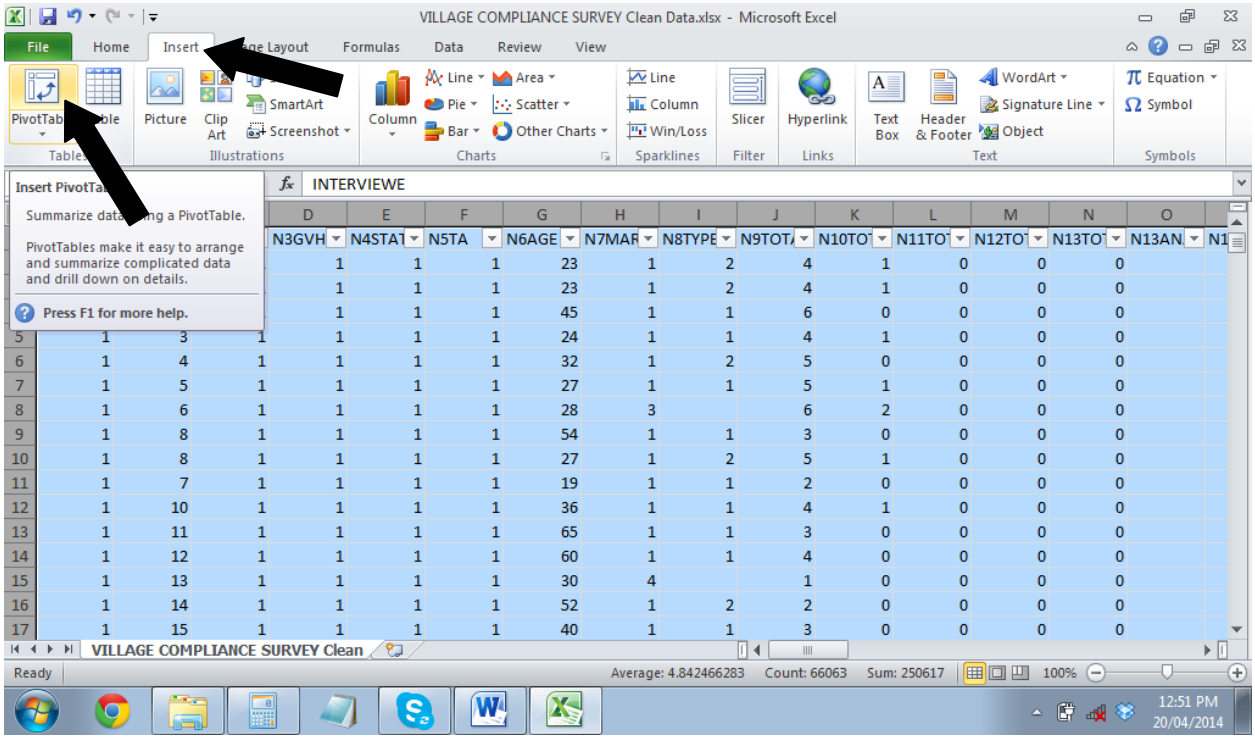
5.1 Creating a Pivot Table

The first step in data analysis is to create a pivot table. You can then use the pivot table to create tables showing the results that you want.

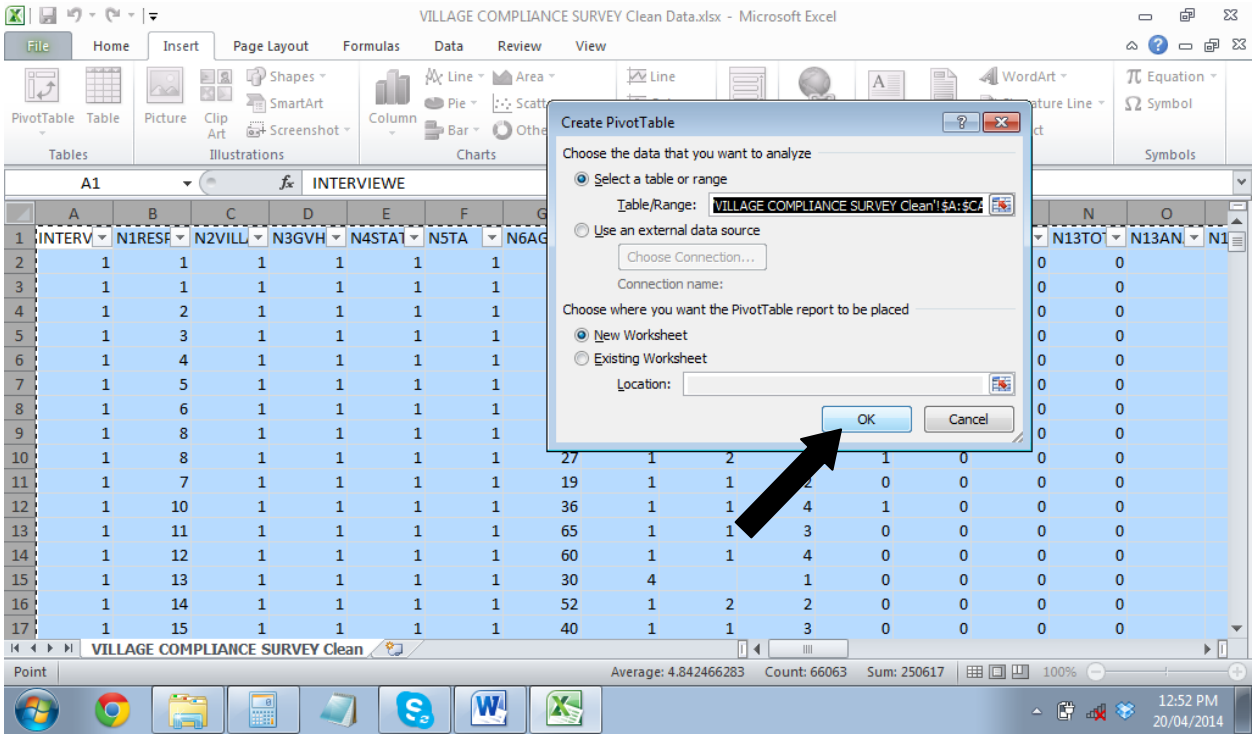
Step 1: Select all the rows and columns in the spreadsheet by clicking in the very top left corner. Or press the Ctrl + A keys on the keyboard.



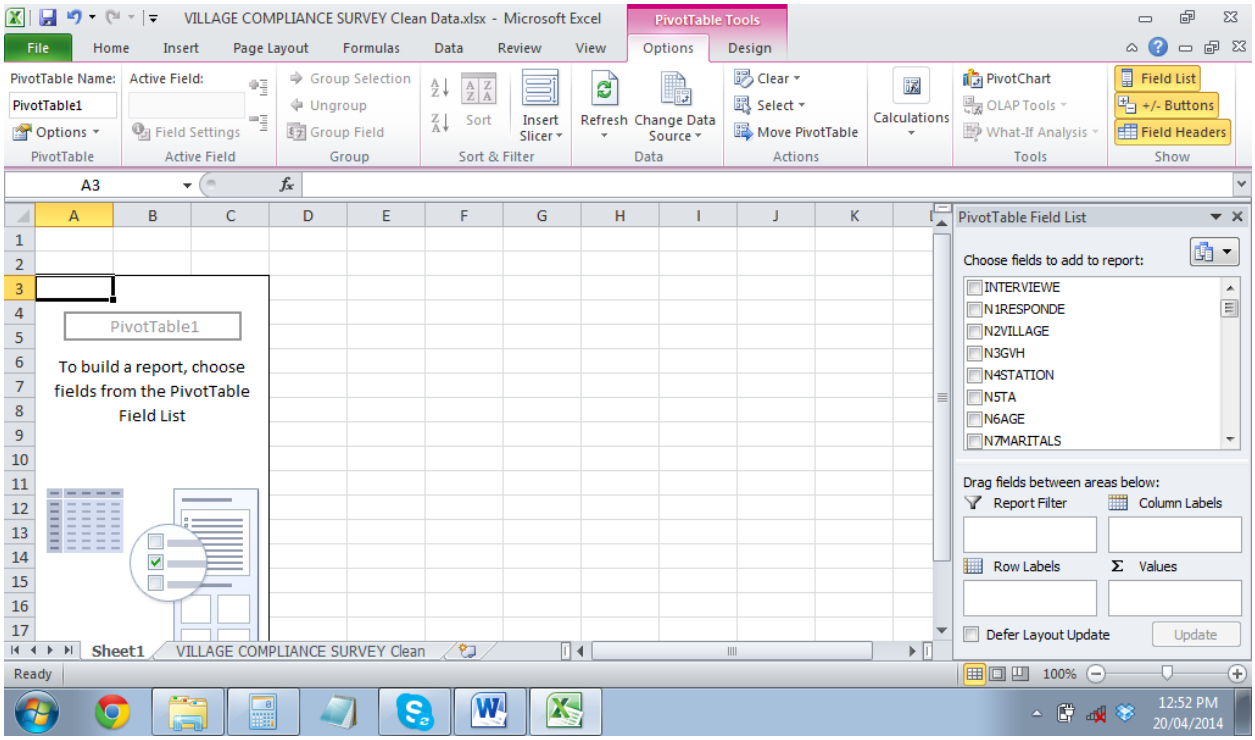
Step 2: Click on the Insert menu item, then click on the Pivot Table button.



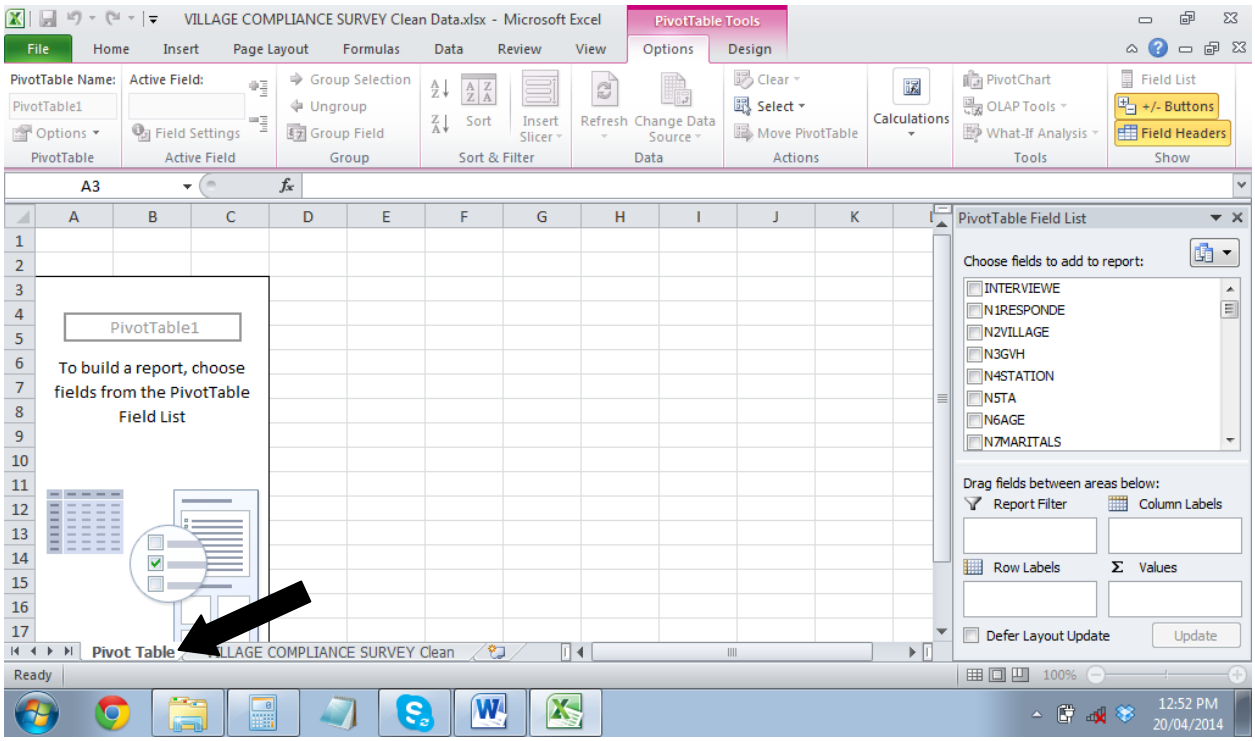
Step 3: Leave all the settings as they are, just click OK.



Step 4: A new tab will appear in Excel showing the pivot table.



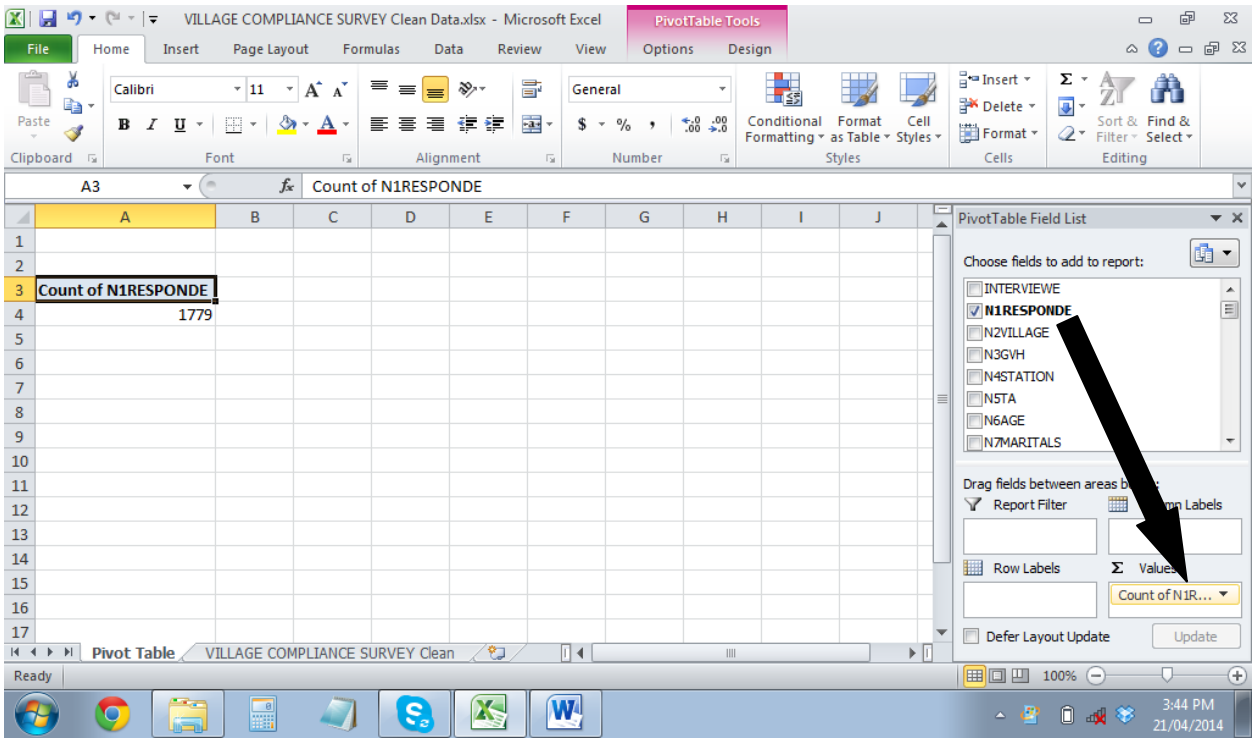
Step 5: If you want to you can change the name of the new tab from Sheet 1 to Pivot Table. To change the name double click on it and type a new one.



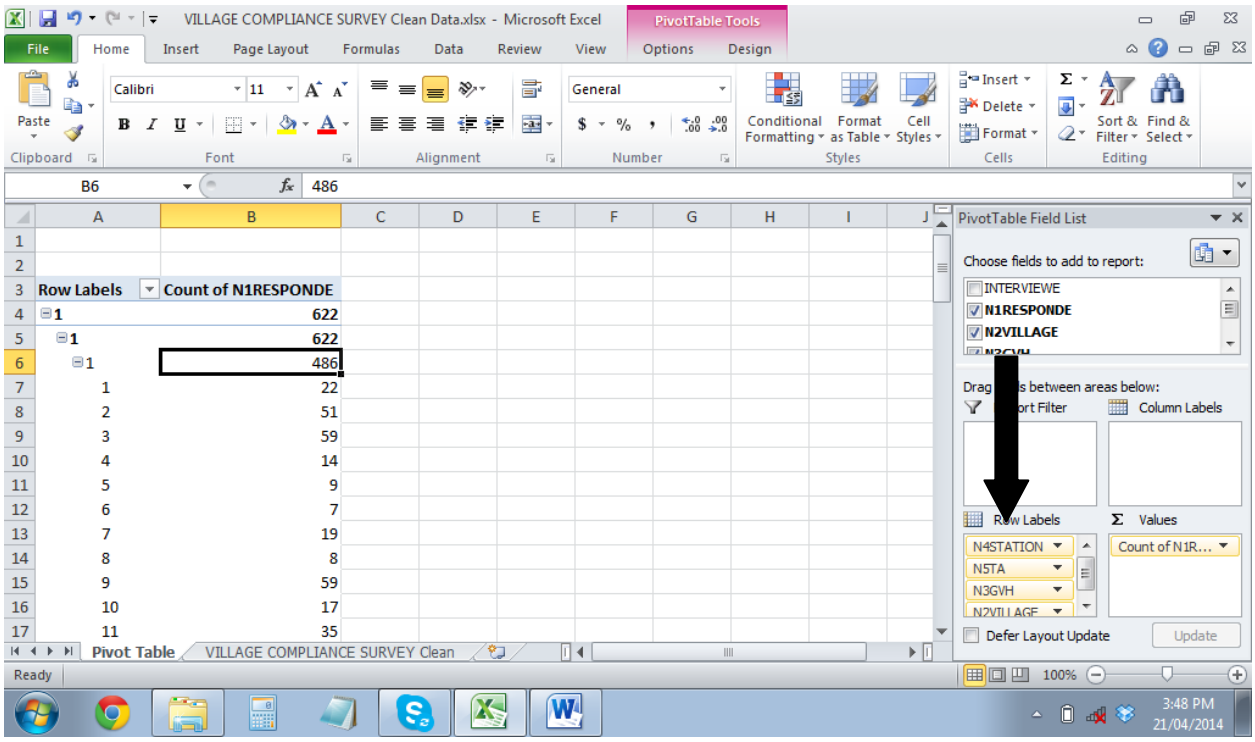
5.2 Checking the codes

Before starting the analysis you need to check how many surveys were completed for each code (station, GVH, village), and that none of the codes have an error.

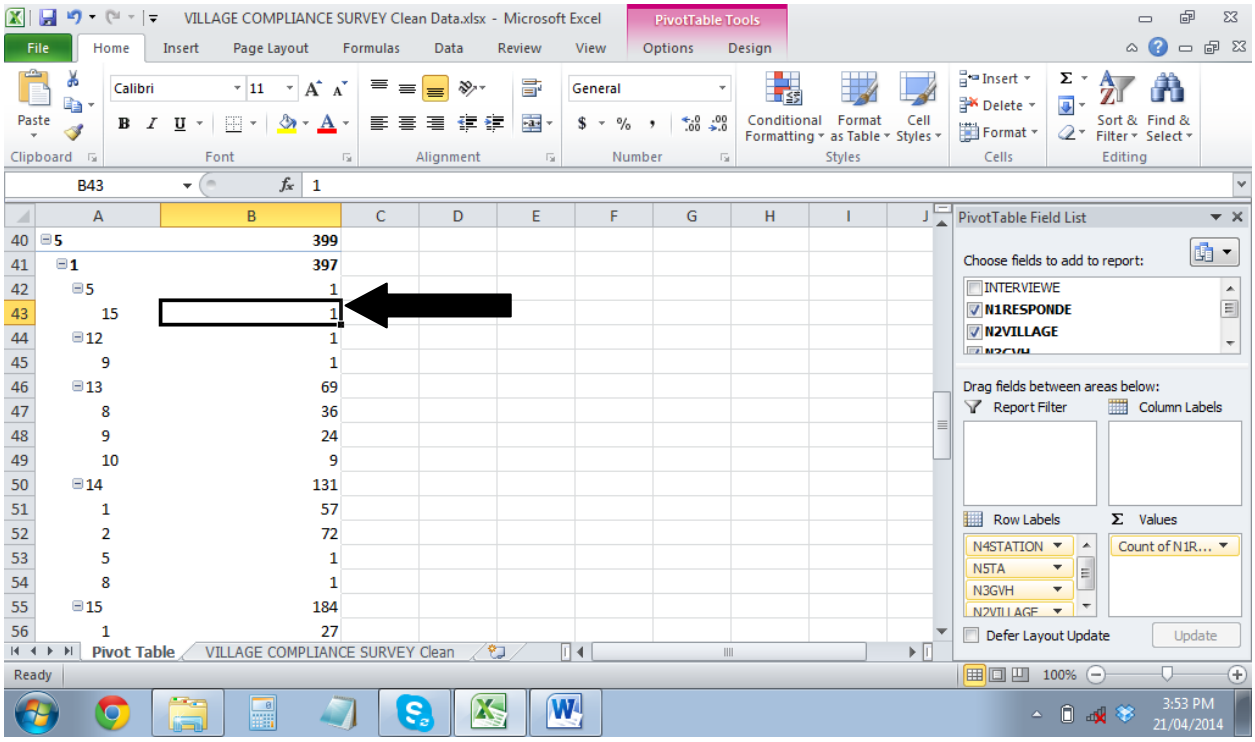
Step 1: On the pivot table screen drag the respondent code (Q1) into the Values box. Make sure it says “Count of…”. In the pivot table you will see a number. This is the total number of surveys.



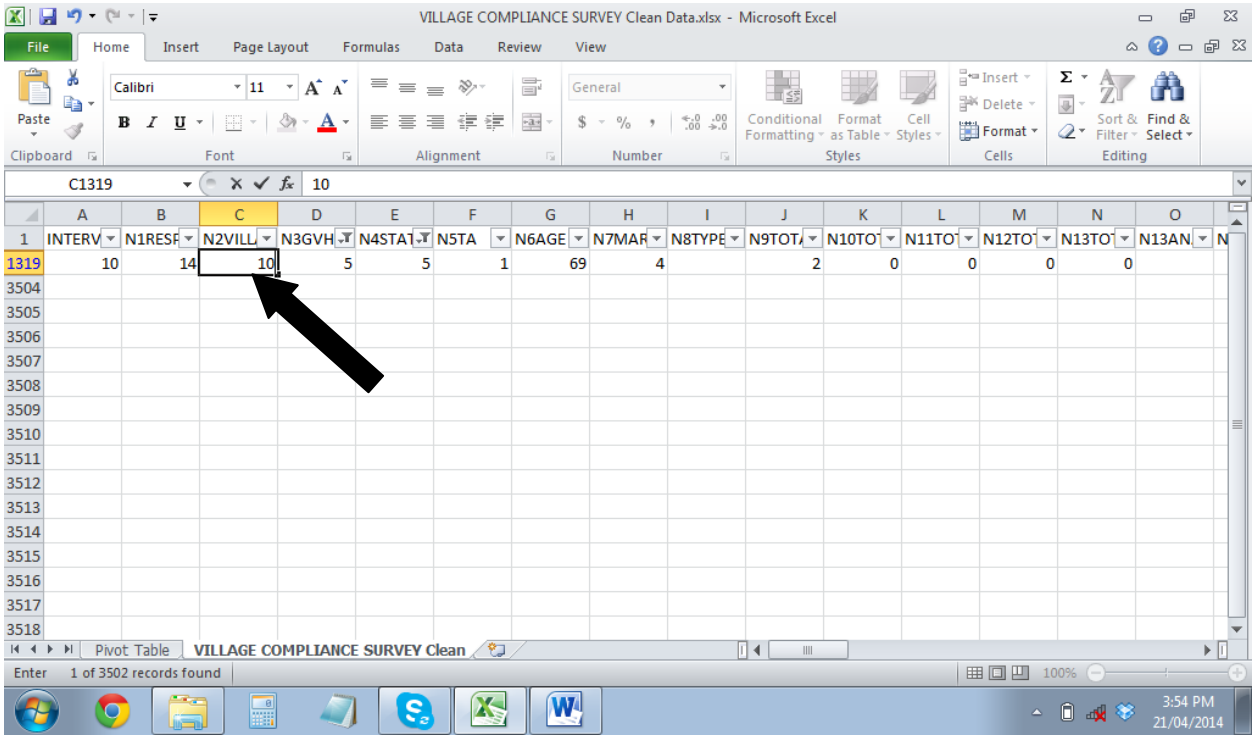
Step 2: Drag the station code (Q4), TA code (Q5), GVH code (Q3) and village code (Q2) into the Row Labels box, in that order. Station code should be at the top of the box, followed by TA, GVH, then village. In the pivot table you will see the code numbers on the left side, with the total number of surveys on the right side. The top level of codes is the station, followed by the TA, the GVH and then the village.



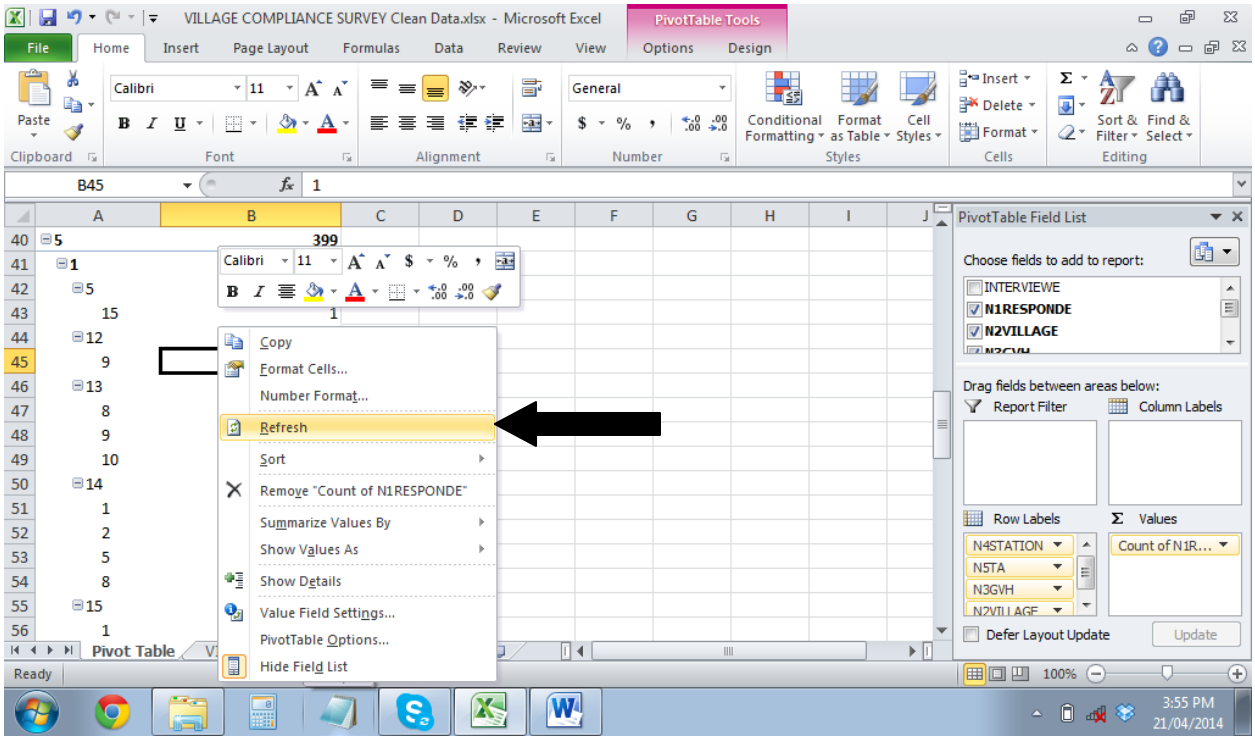
Step 3: Check all the codes to make sure they are in the right place compared to the code list. Look out for village codes that only have 1 survey – these are likely to be errors. In the example below station #5 (Nyamazani) has a village #15 under GVH #5 that only has one survey. According to the code list station #5 should not have any villages under GVH #5. This is an error.



Step 4: Fix the error by going back to the data sheet. Use the filters to find the survey with the error (in this example filter station #5, TA #1, GVH #5, village #15). Find the original hard copy of the survey and correct the codes on the spreadsheet.



Step 5: Go back to the Pivot Table tab. Right click anywhere on the pivot table and click Refresh. The table should update and the survey with the error will disappear. You can now proceed with the next code correction.

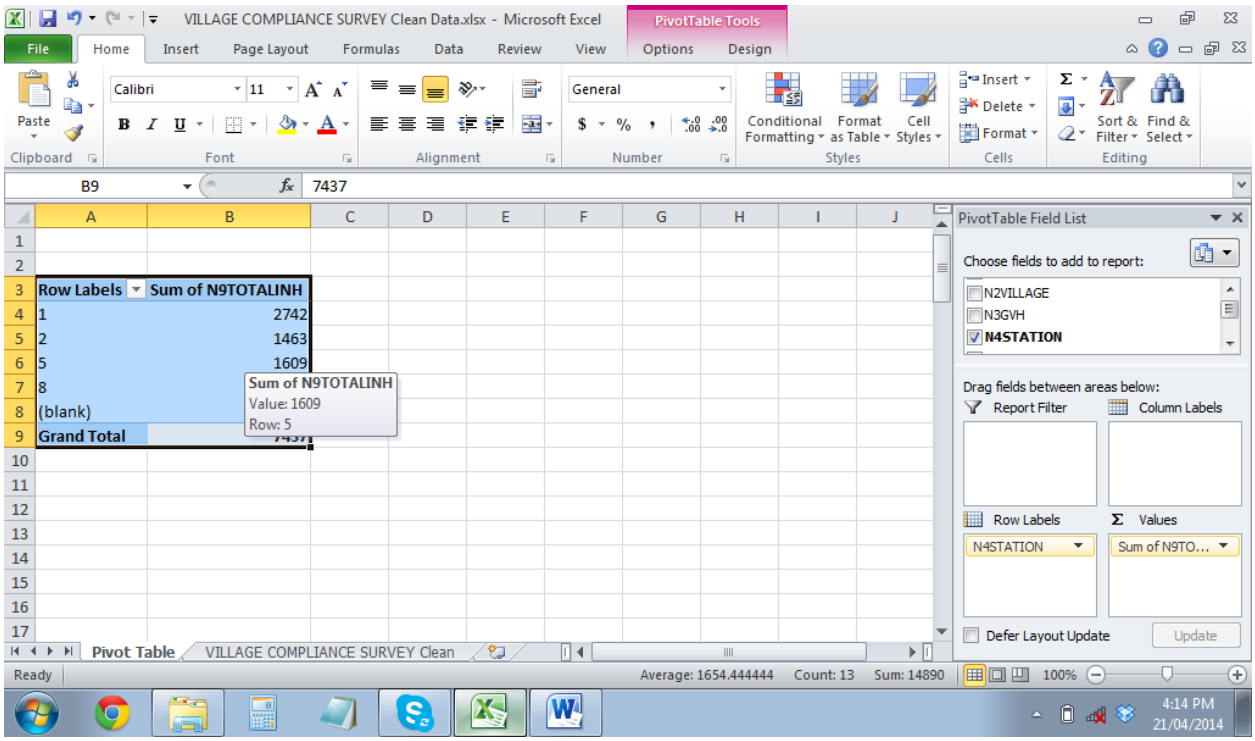


5.3 Copying and pasting from Pivot Tables

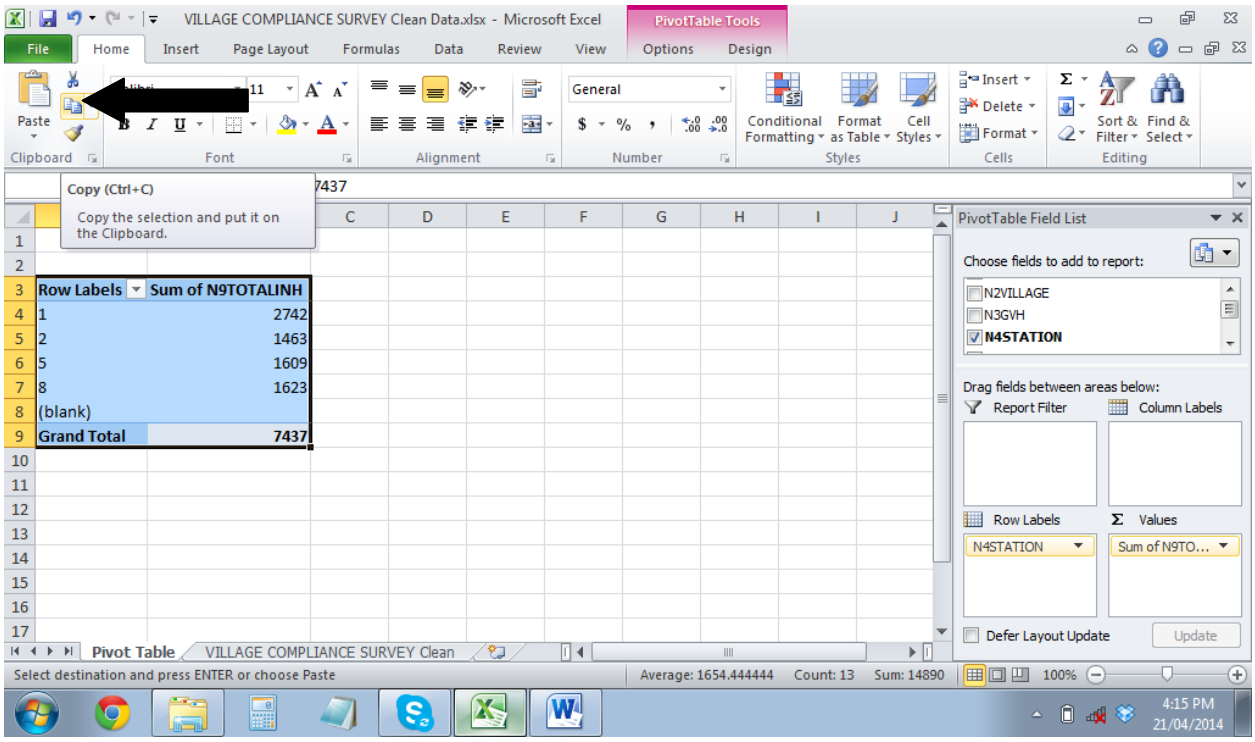
Throughout the analysis you will need to copy and paste Pivot Tables to Word and Excel.

5.3.1 Copying to Word

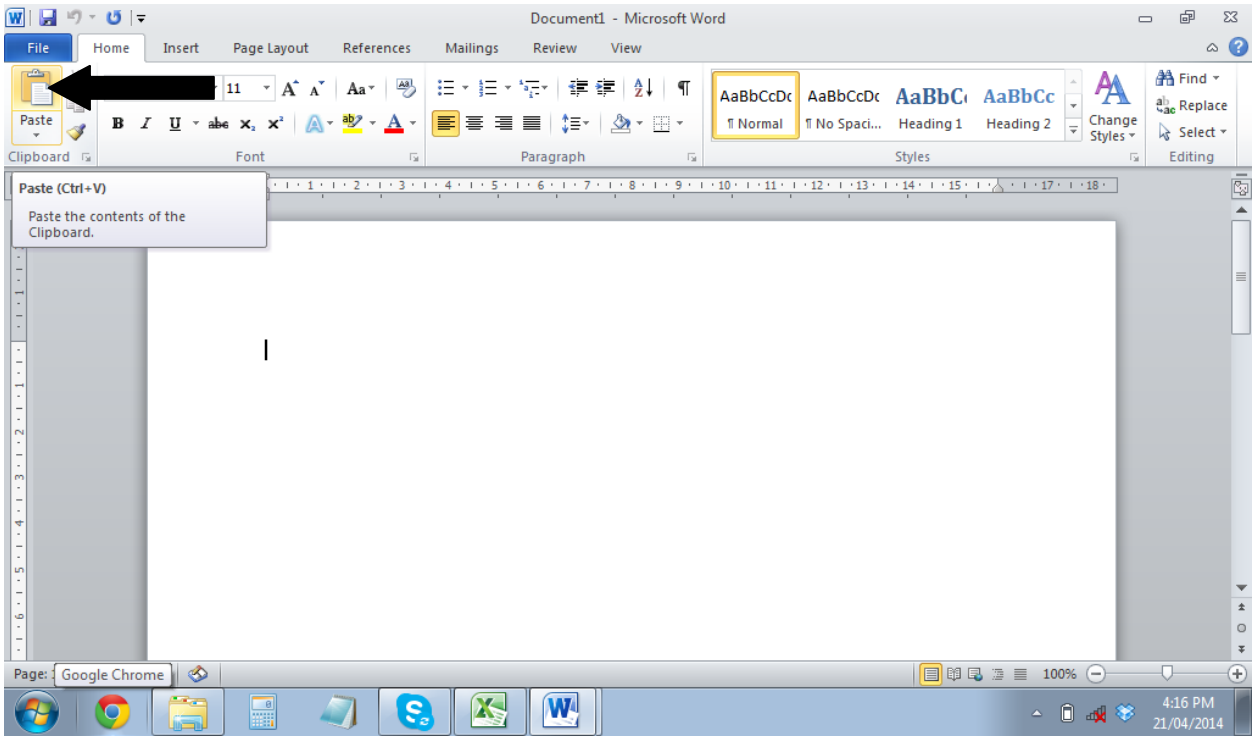
Step 1: Select the pivot table.



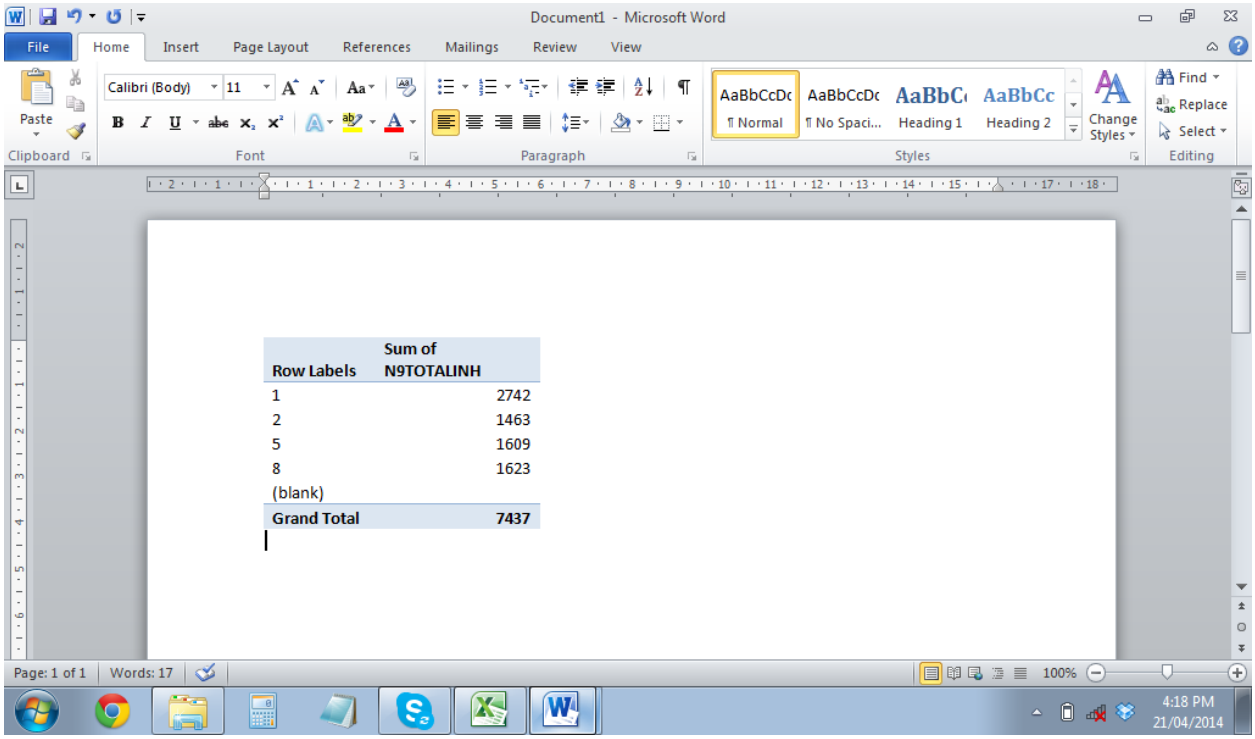
Step 2: Press the copy button in the top left of the screen (or press the Ctrl + C buttons on the keyboard).



Step 3: Open the Word document and press the Paste button (or Ctrl + P on the keyboard).

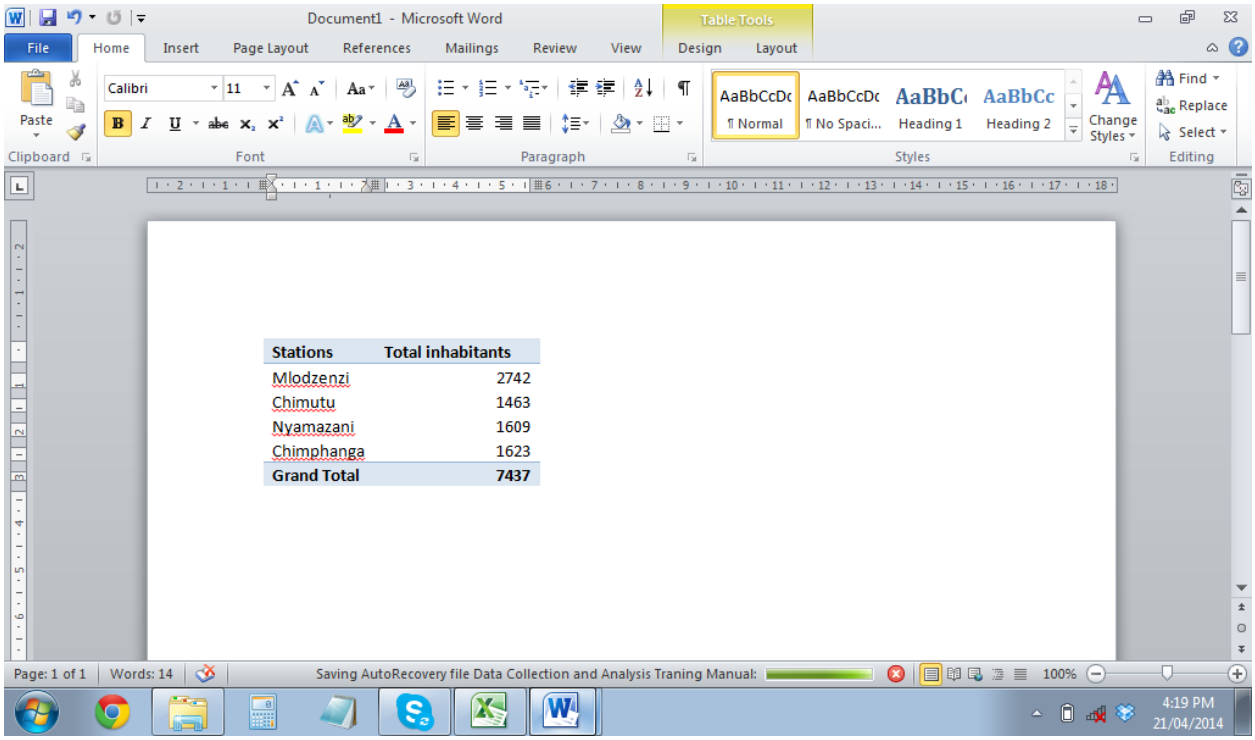


Step 4: The table will appear.



Row Labels	Sum of N9TOTALINH
1	2742
2	1463
5	1609
8	1623
(blank)	
Grand Total	**7437**

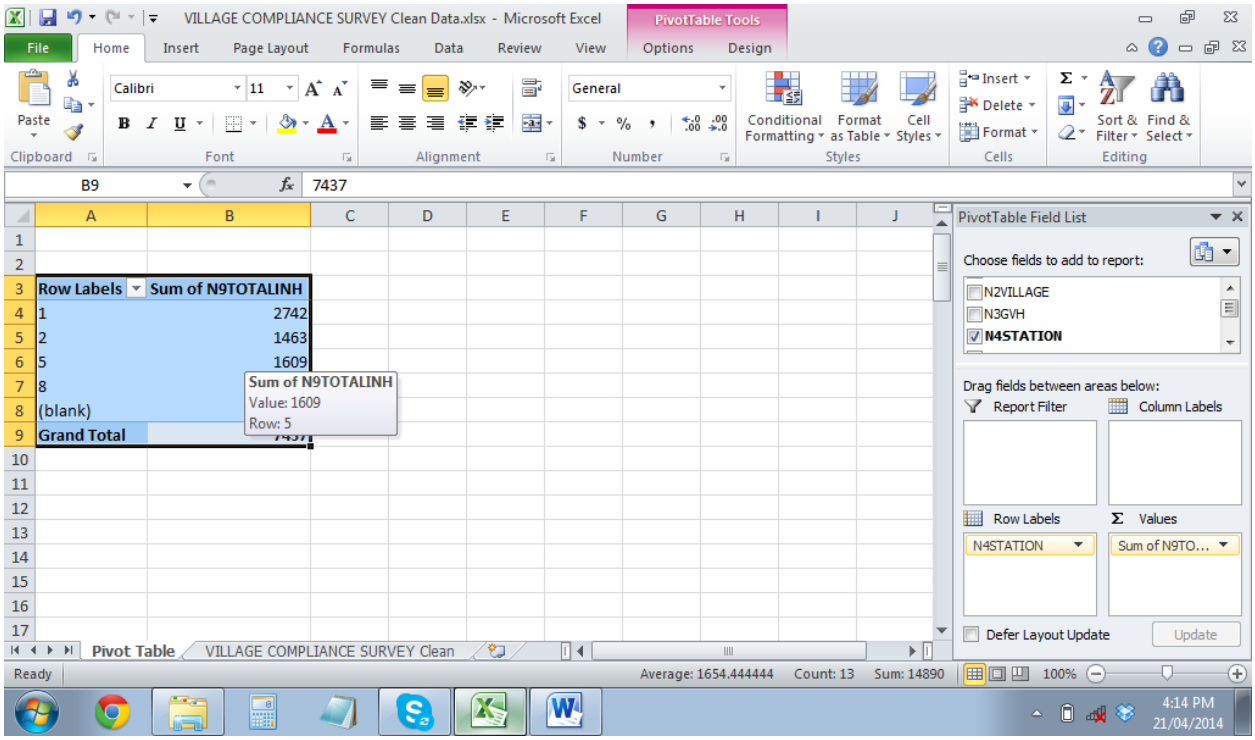
Step 5: Replace the code numbers with written names and change the column headings to something sensible. Delete any unnecessary rows, such as those with blank entries.



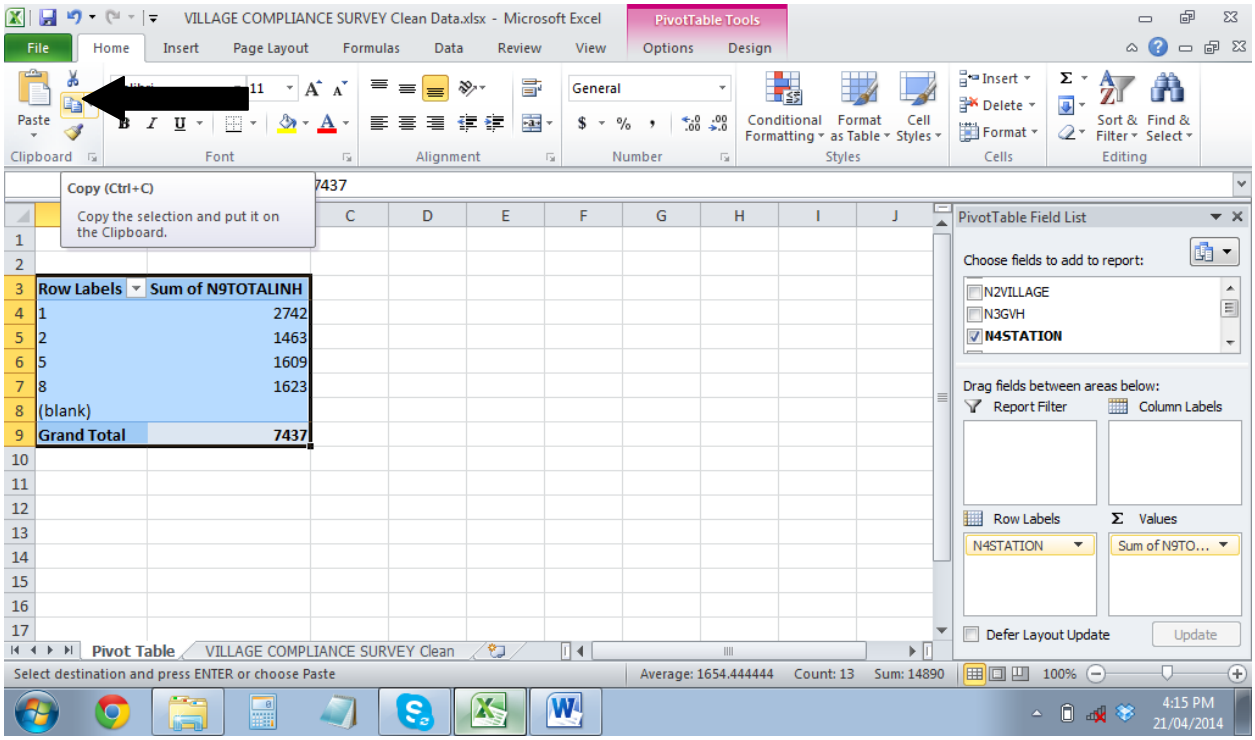
Stations	Total inhabitants
Mlodzenzi	2742
Chimutu	1463
Nyamazani	1609
Chimphanga	1623
Grand Total	**7437**

5.3.2 Copying to Excel

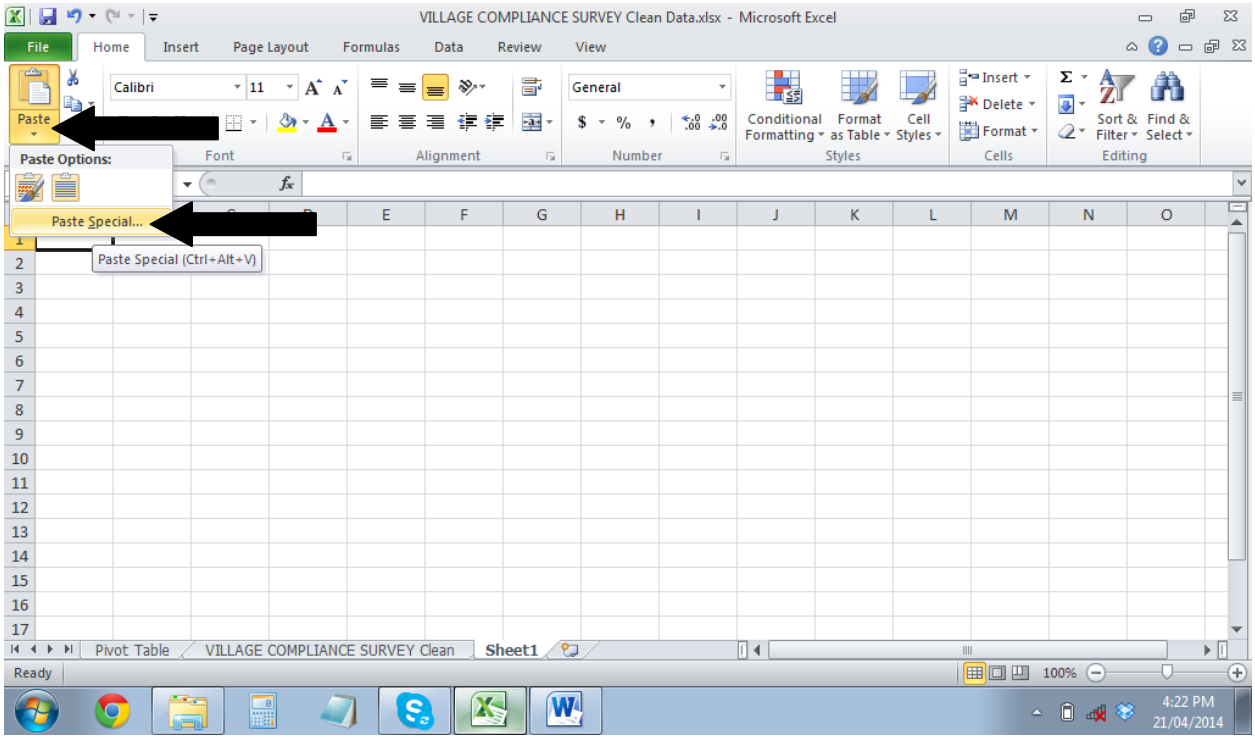
Step 1: Select the pivot table.



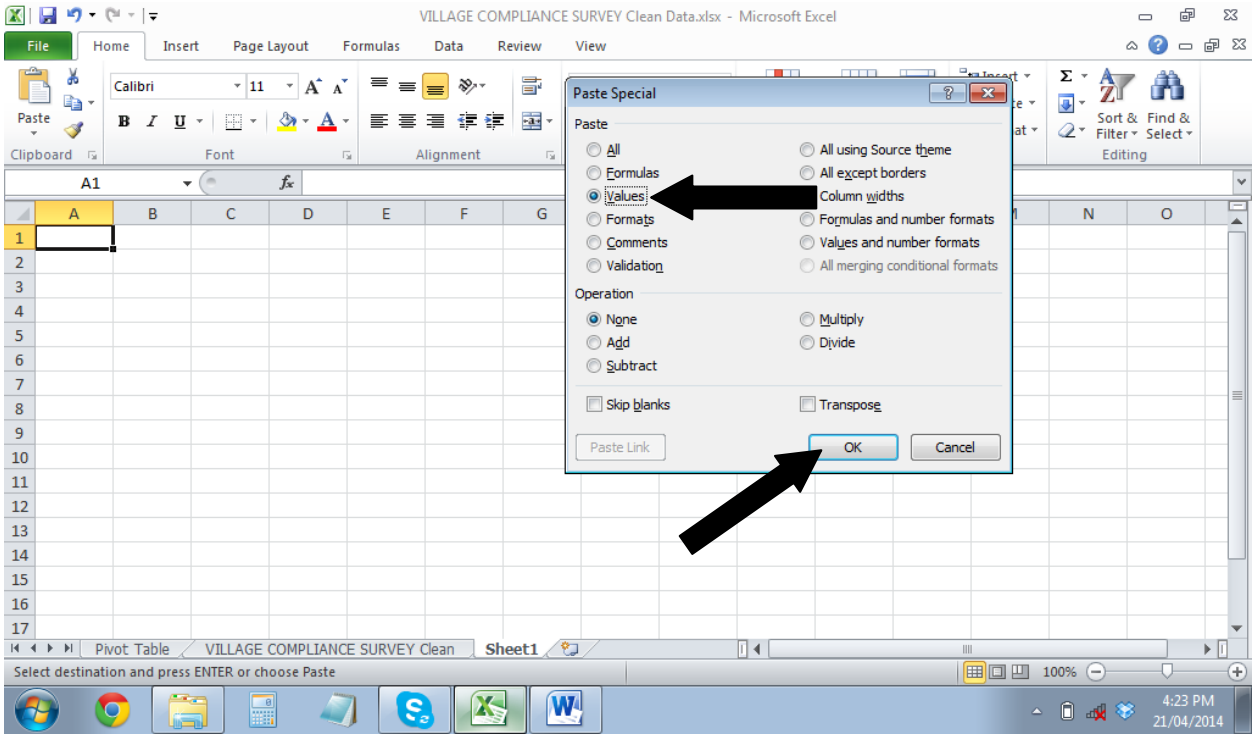
Step 2: Press the copy button in the top left of the screen (or press the Ctrl + C buttons on the keyboard).



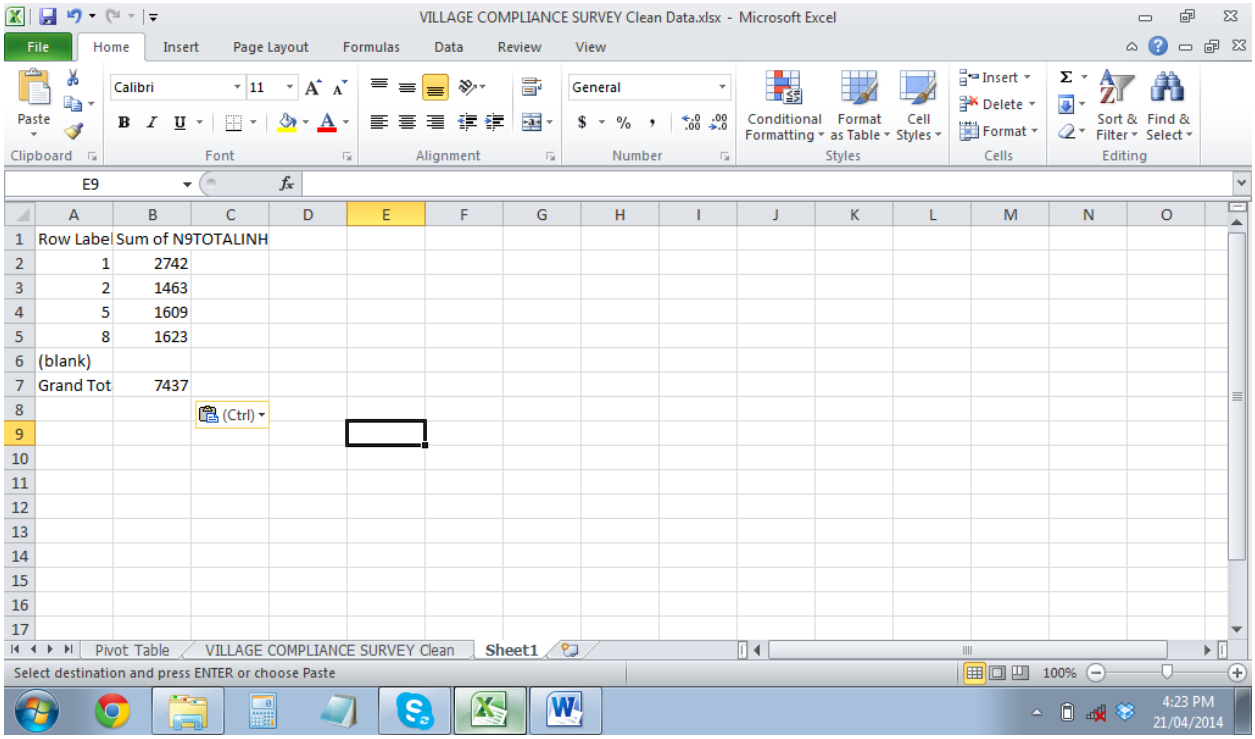
Step 3: Open the spreadsheet where you want to paste the table. Click on the arrow under the Paste button. Choose Paste Special.



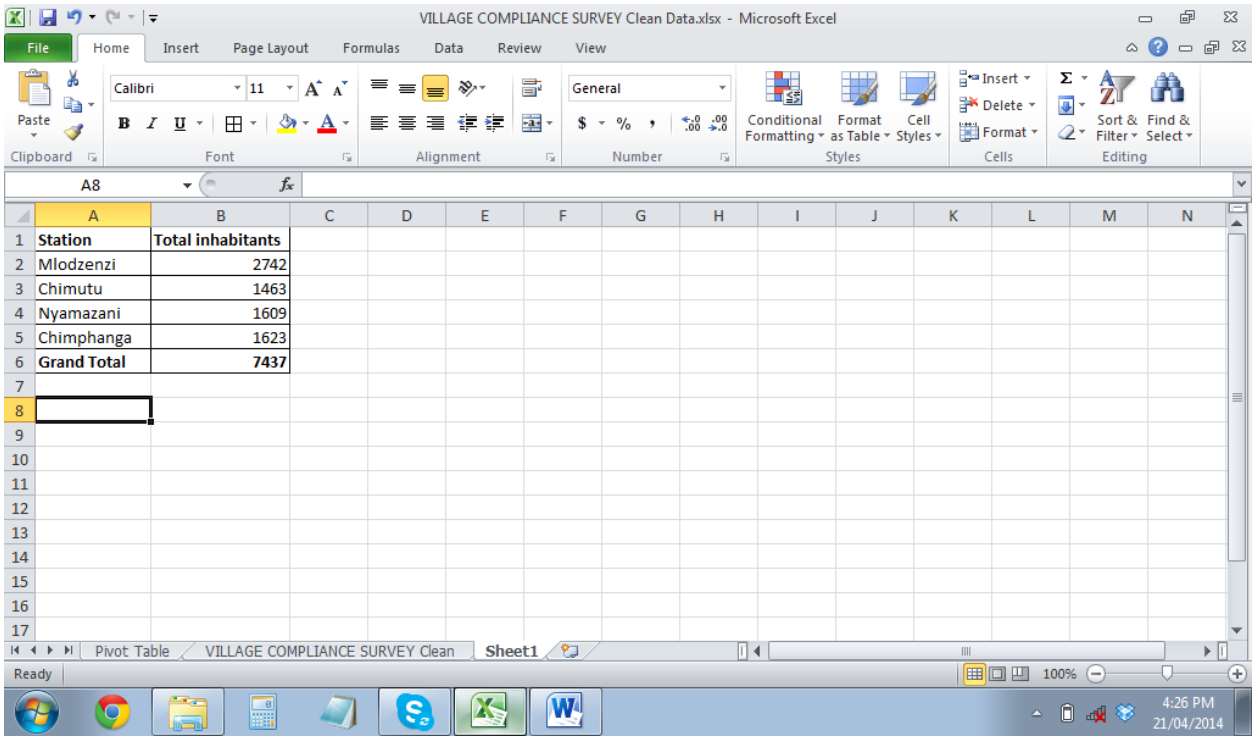
Step 4: Change the option from All to Values. Click OK.



Step 5: The table will appear.



Step 6: Replace the code numbers with written names and change the column headings to something sensible. Delete any unnecessary rows, such as those with blank entries. Apply formatting (bold, table lines, etc) if necessary.



5.4 Questions where the answer is a number

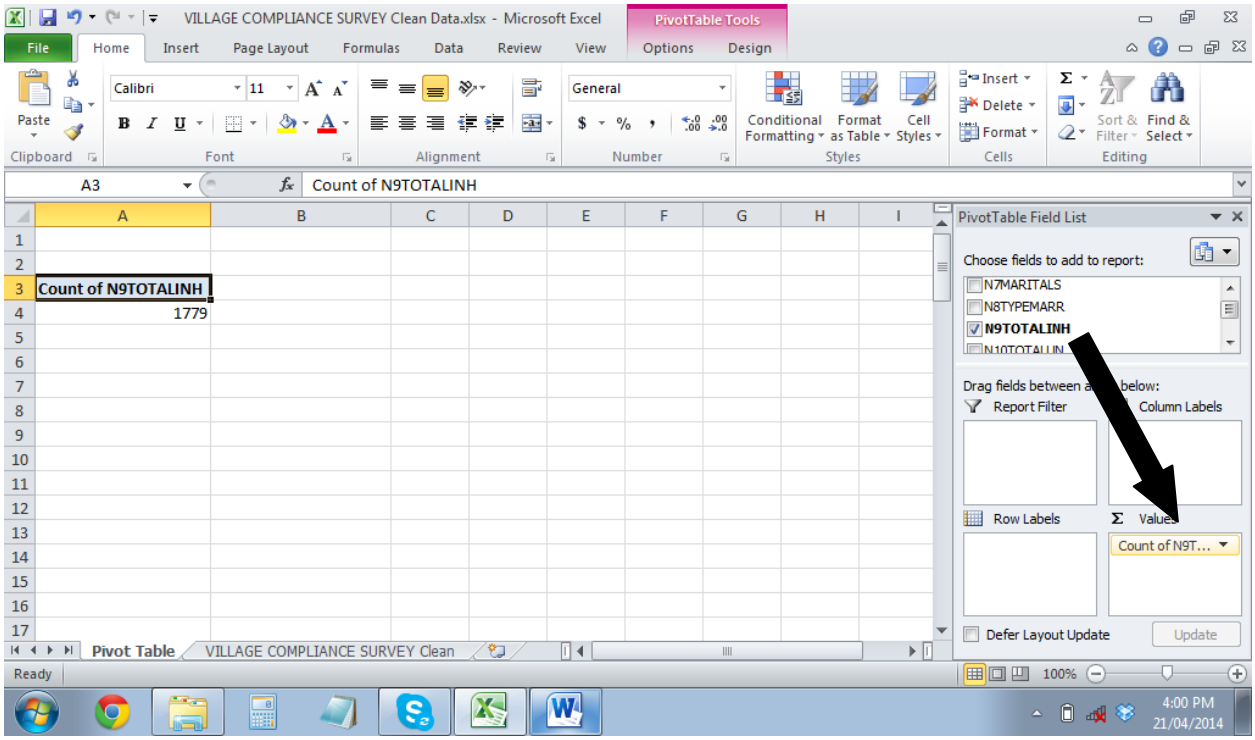
The answer for these questions is always written on a _____.

Examples

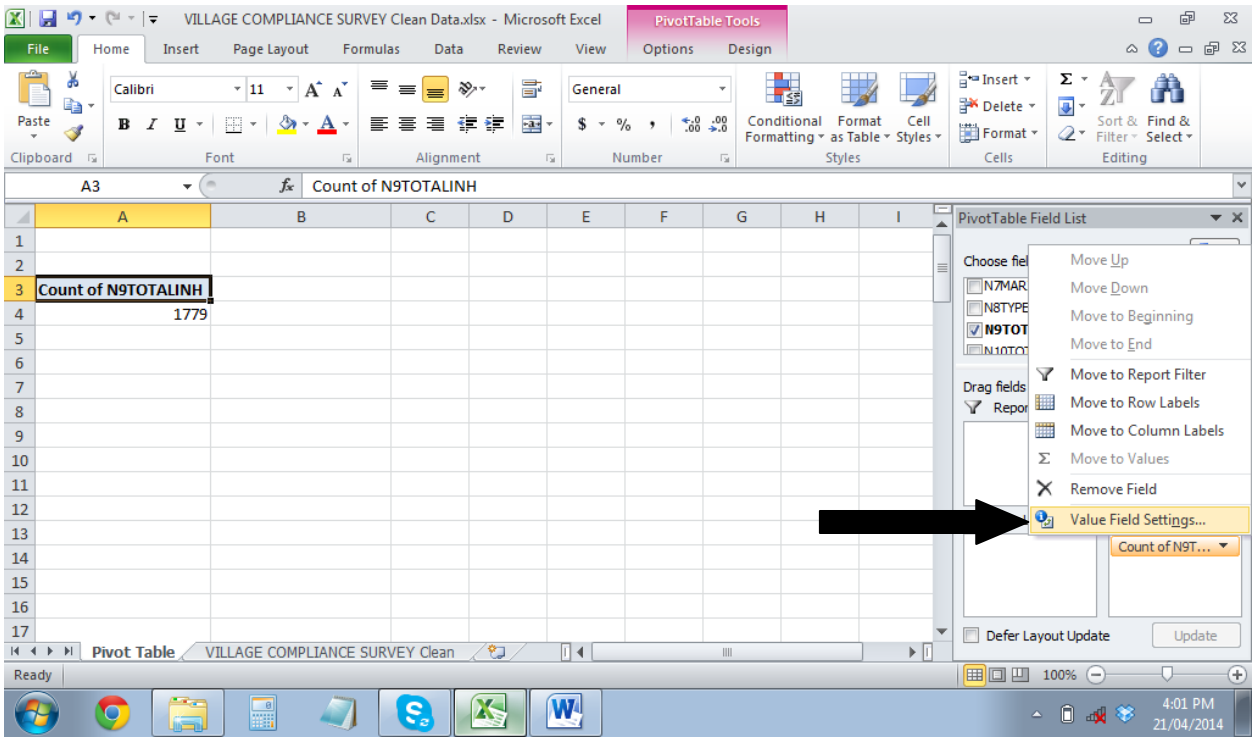
Q6. Age of respondent _____ *Q9. Total inhabitants* _____

5.4.1 Calculating totals

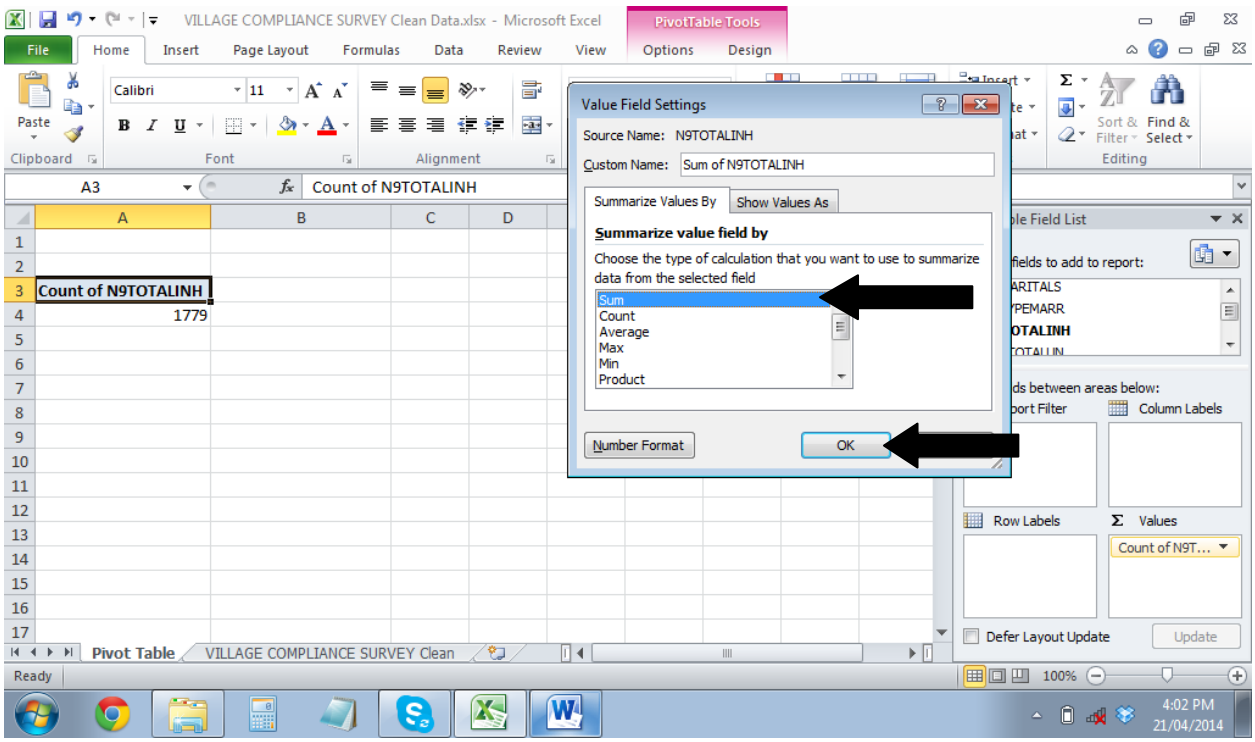
Step 1: Drag the question number (e.g. Q9. Total inhabitants) into the Values box.



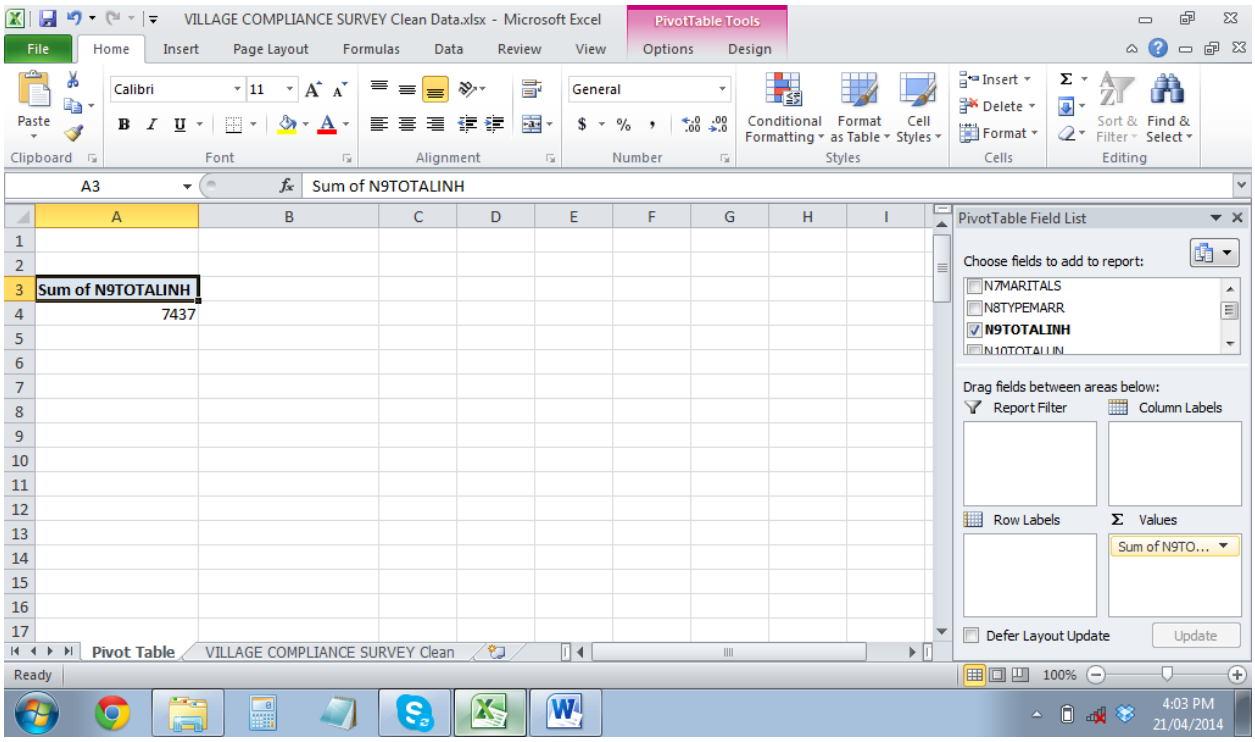
Step 2: Click on the small arrow next to the question in the Values box. Choose Value Field Settings.



Step 3: Change from Count to Sum. Then click OK.



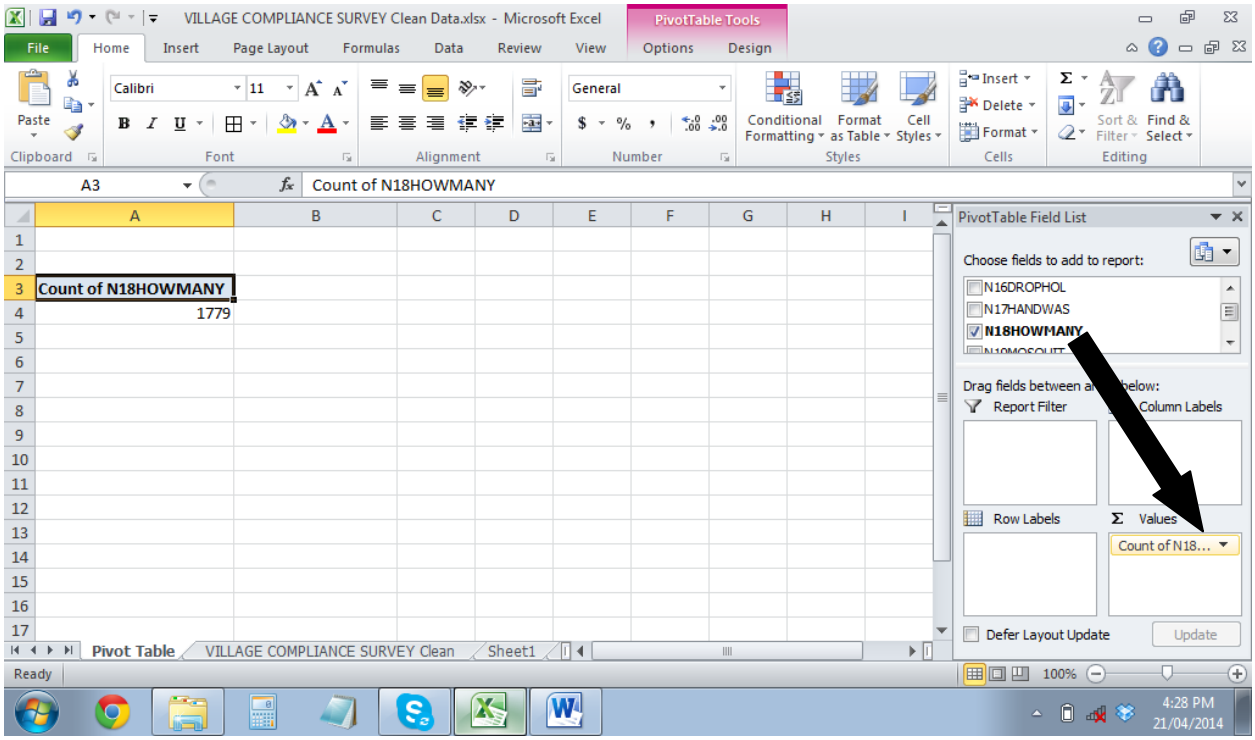
Step 4: The pivot table now shows the total for the question. In this case, it shows the total inhabitants for all households.



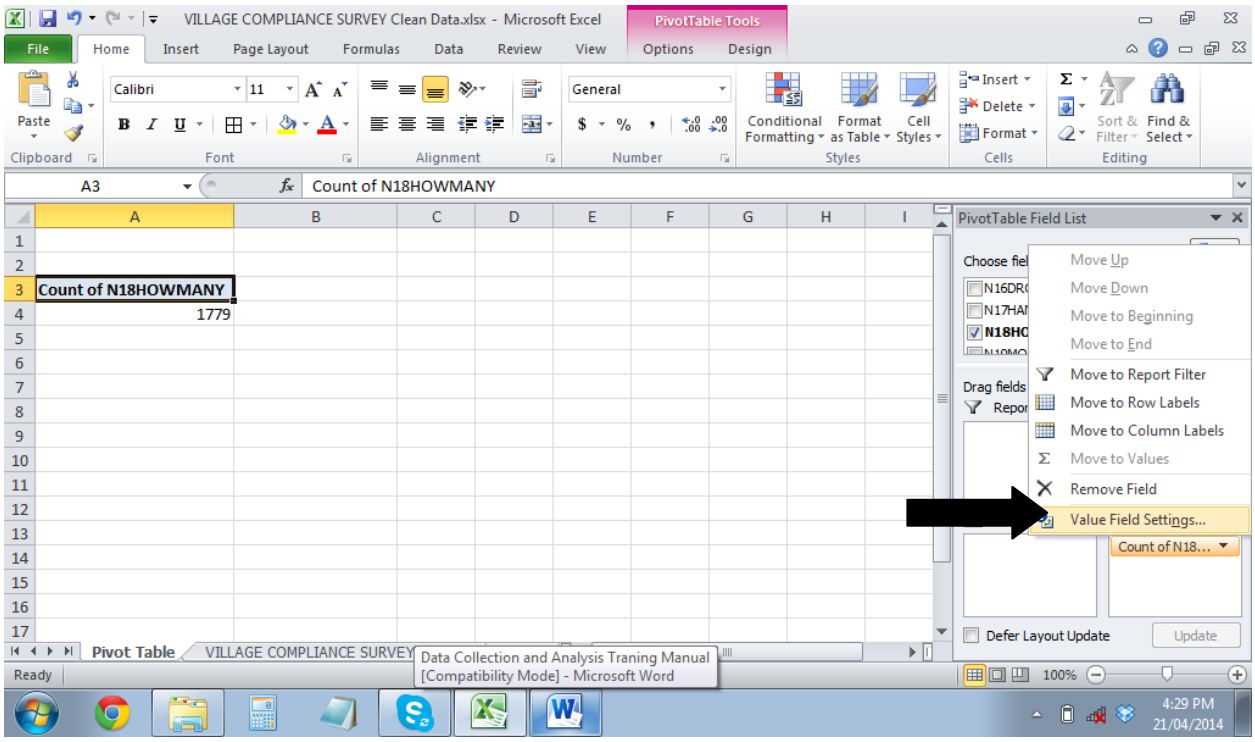
To see the results by station, TA, GVH or Village follow the instructions in section 5.8

5.4.2 Calculating averages

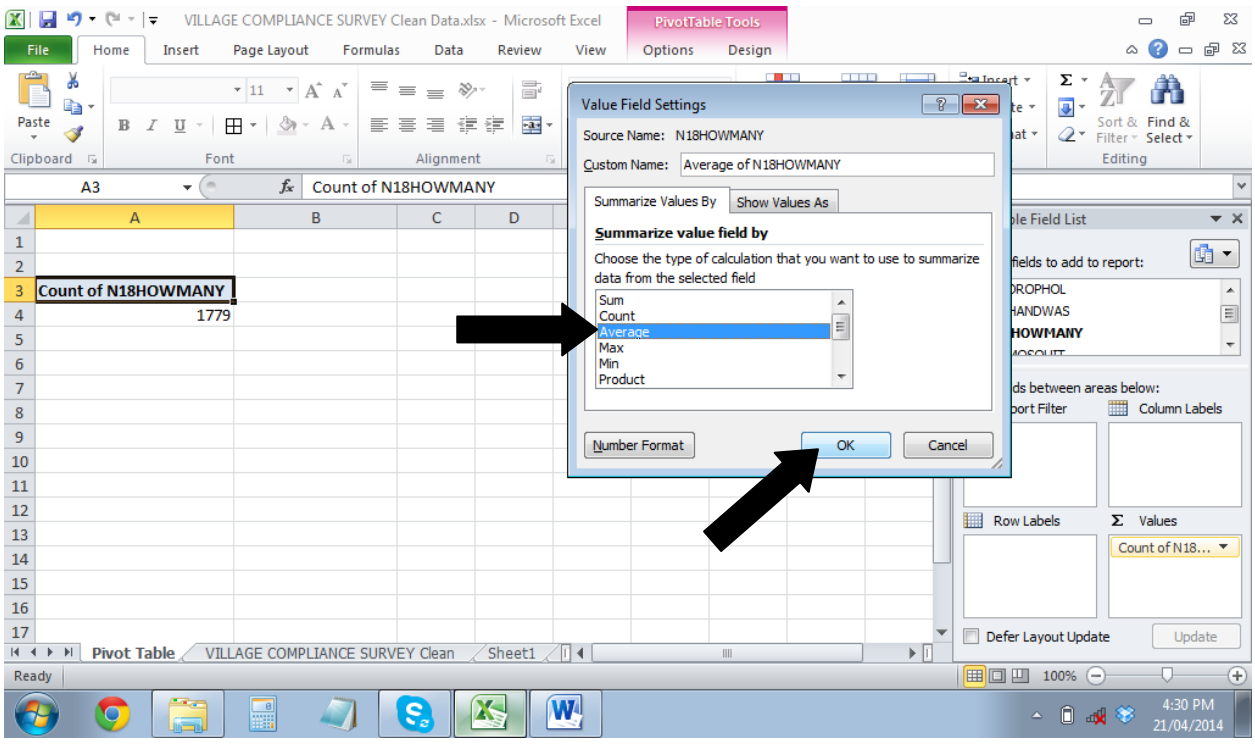
Step 1: Drag the question number (e.g. Q18. How many rooms are used for sleeping?) into the Values box.



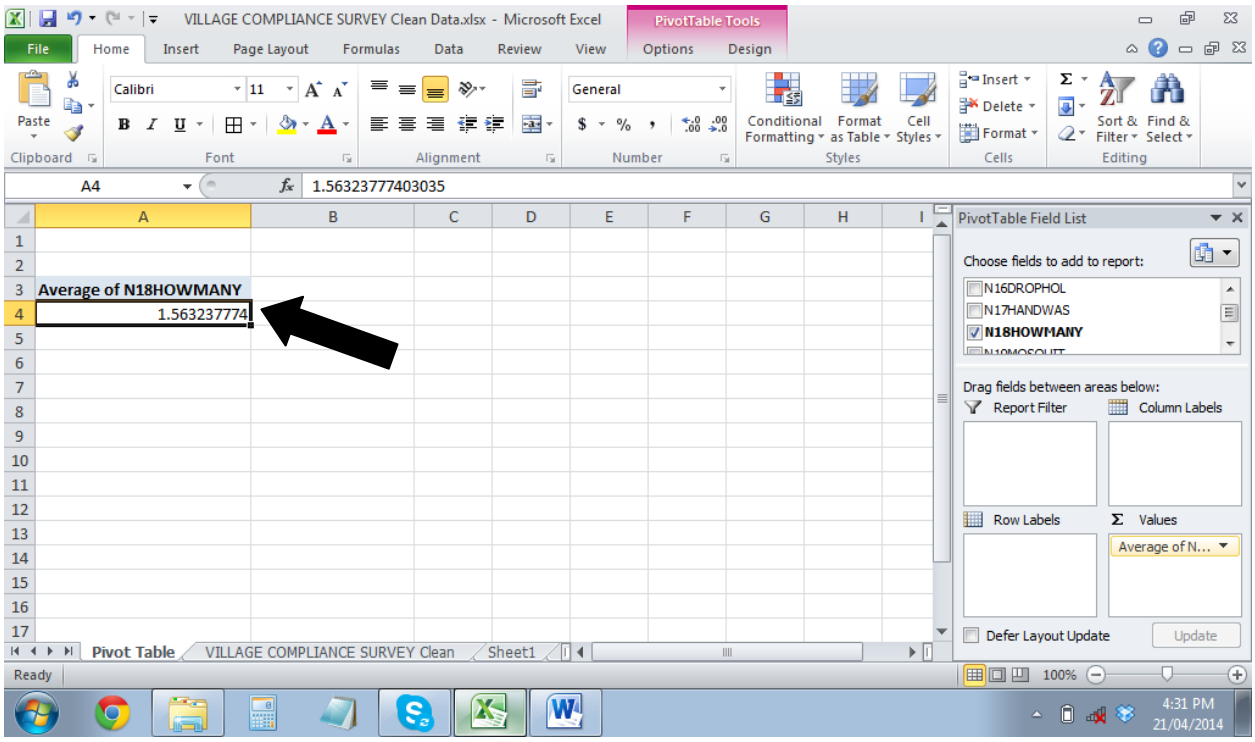
Step 2: Click on the small arrow next to the question in the Values box. Choose Value Field Settings.



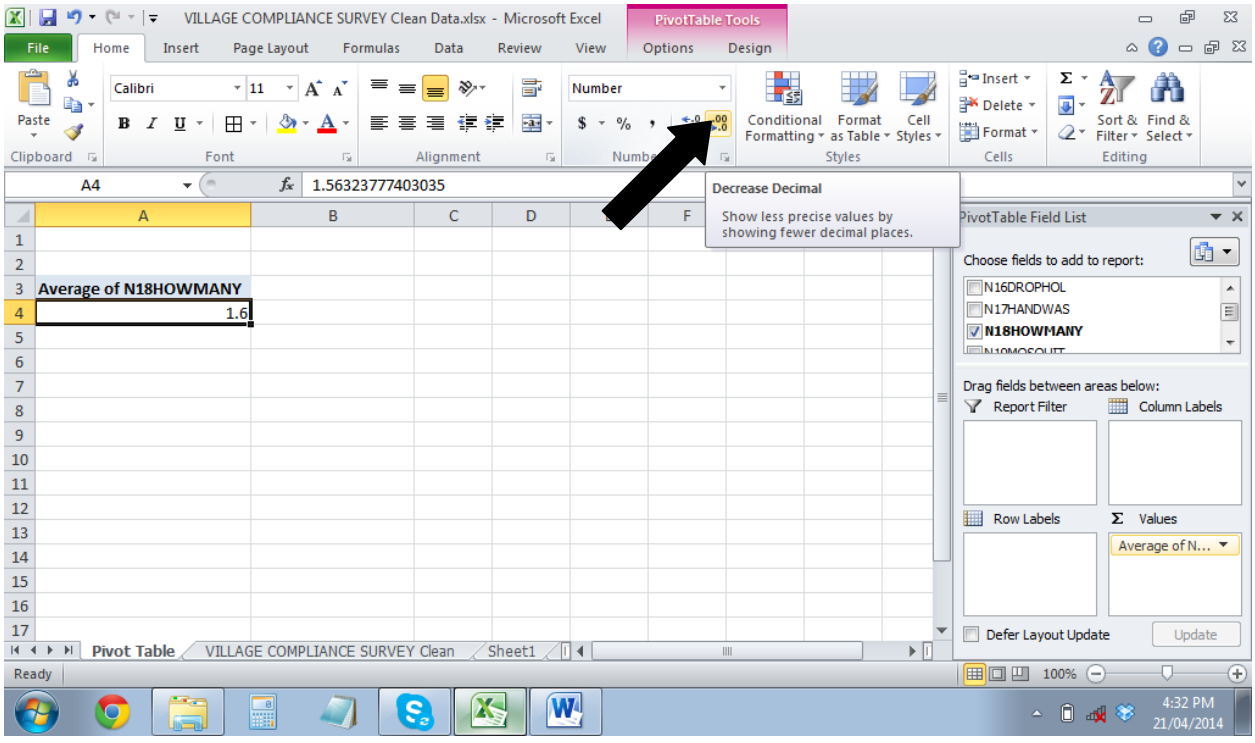
Step 3: Click on the small arrow next to the question in the Values box. Choose Value Field Settings.



Step 4: The pivot table will now show the average. In this case it is the average number of rooms used for sleeping across all households.



Step 5: To reduce the number of decimal places, select the average. Then press on the button shown in the following image to remove one decimal place. Keep pressing until you have the desired number of decimal places.

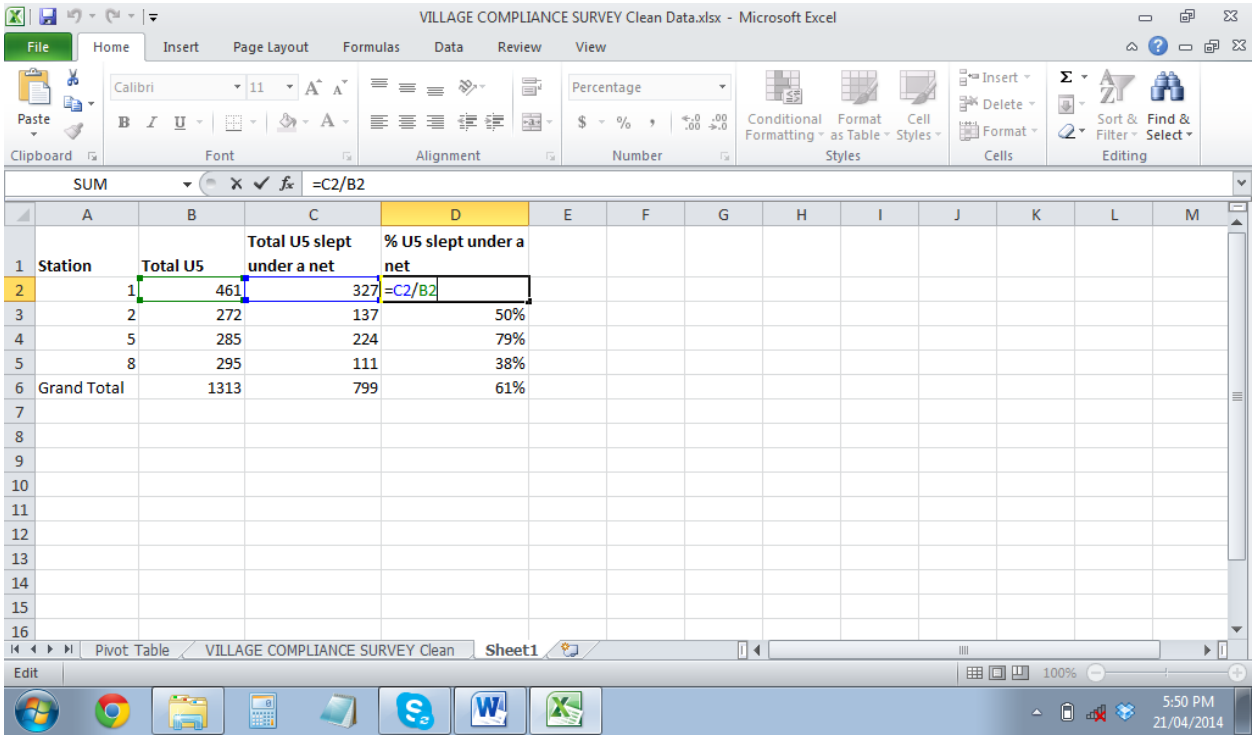


To see the results by station, TA, GVH or Village follow the instructions in section 5.8

5.4.3 Calculating percentages

Sometimes you may want to calculate a percentage using two totals. For example, you could use the total number of under 5's and the total number of under 5's sleeping under a net to calculate the % of under 5's sleeping under a net.

To do this you should follow the steps to calculate the first total (e.g. total under 5s) and copy the table to a spreadsheet. Then follow the steps again to calculate the second total (e.g. total under 5s who slept under a bed net). Copy those results to the spreadsheet too. Then use a formula to calculate the percentage.



5.5 Questions where the answer is multiple choice (tick one)

These questions have a ○ for the answers.

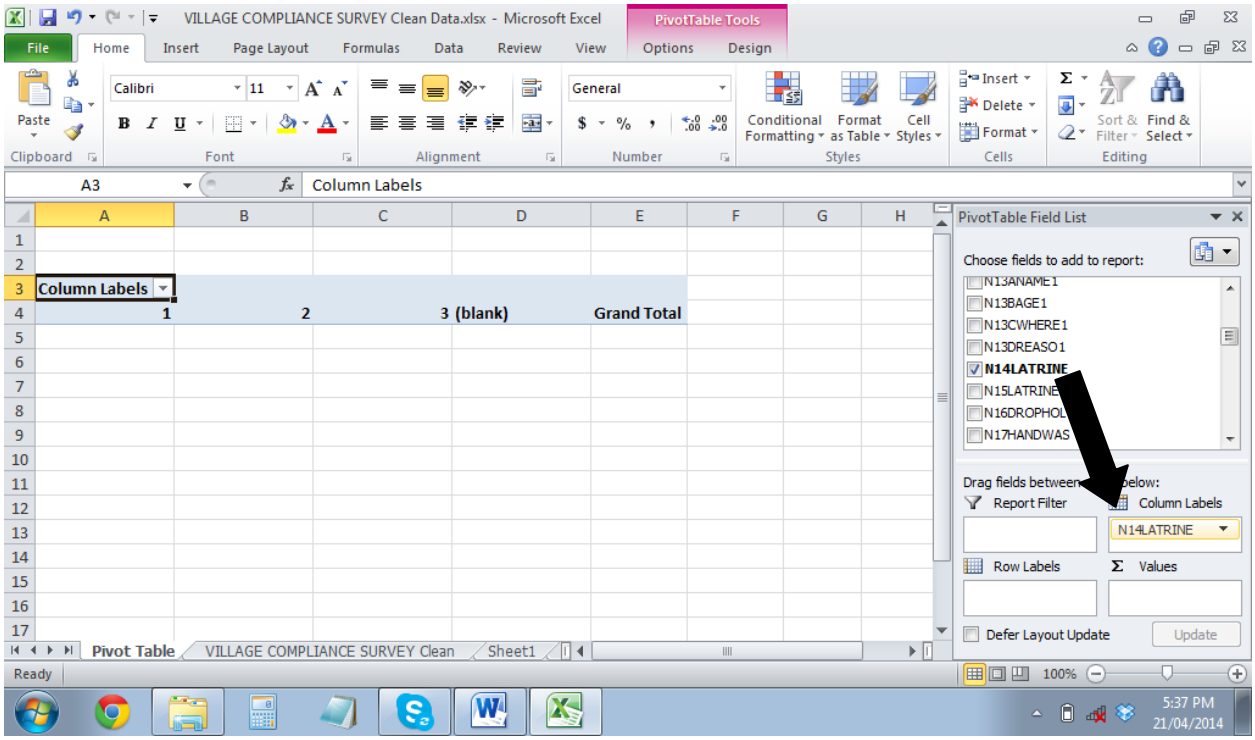
Example

Q19. Does the household have a latrine?

1 ○ Observed Present
2 ○ Observed Not Present
3 ○ No permission to observe

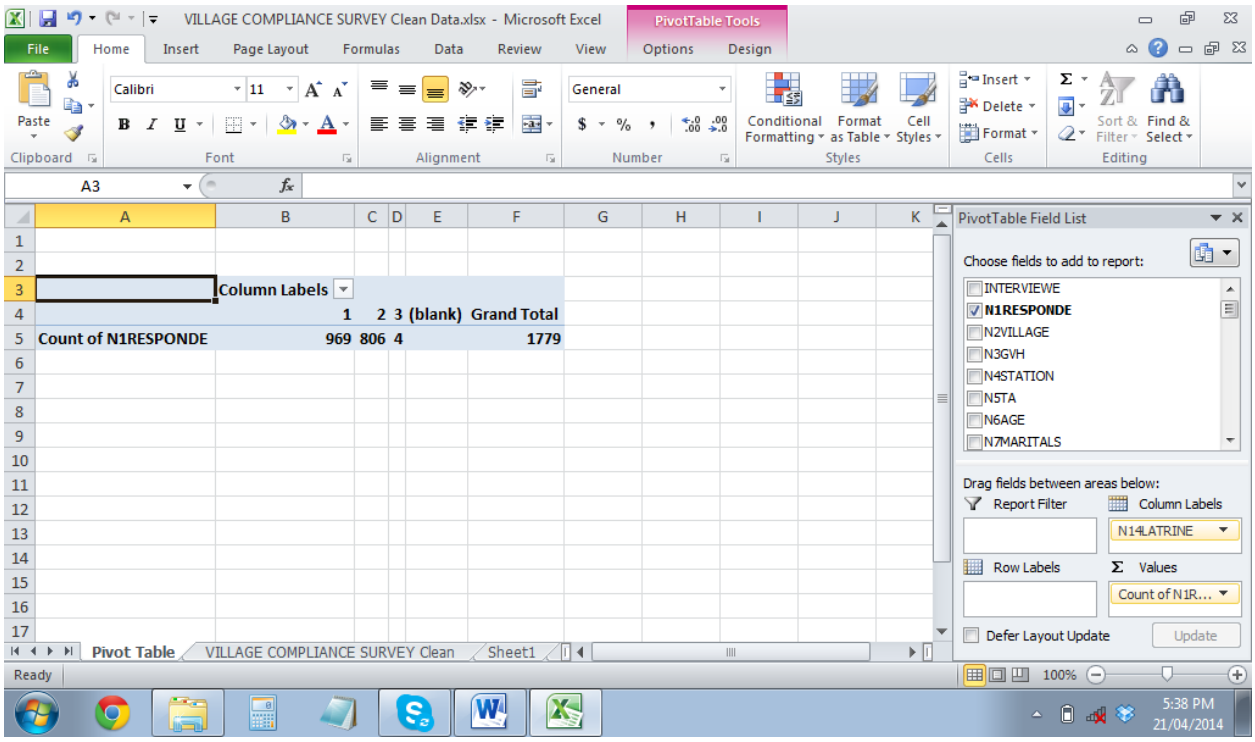
5.5.1 Calculating totals

Step 1: Drag the question (in this case Q19. Does the household have a latrine?) into the Column Labels box.



Step 2: Drag the Respondent Code into the Values box. Make sure it says "Count of…". The pivot table will now show the total number of people who gave each response.

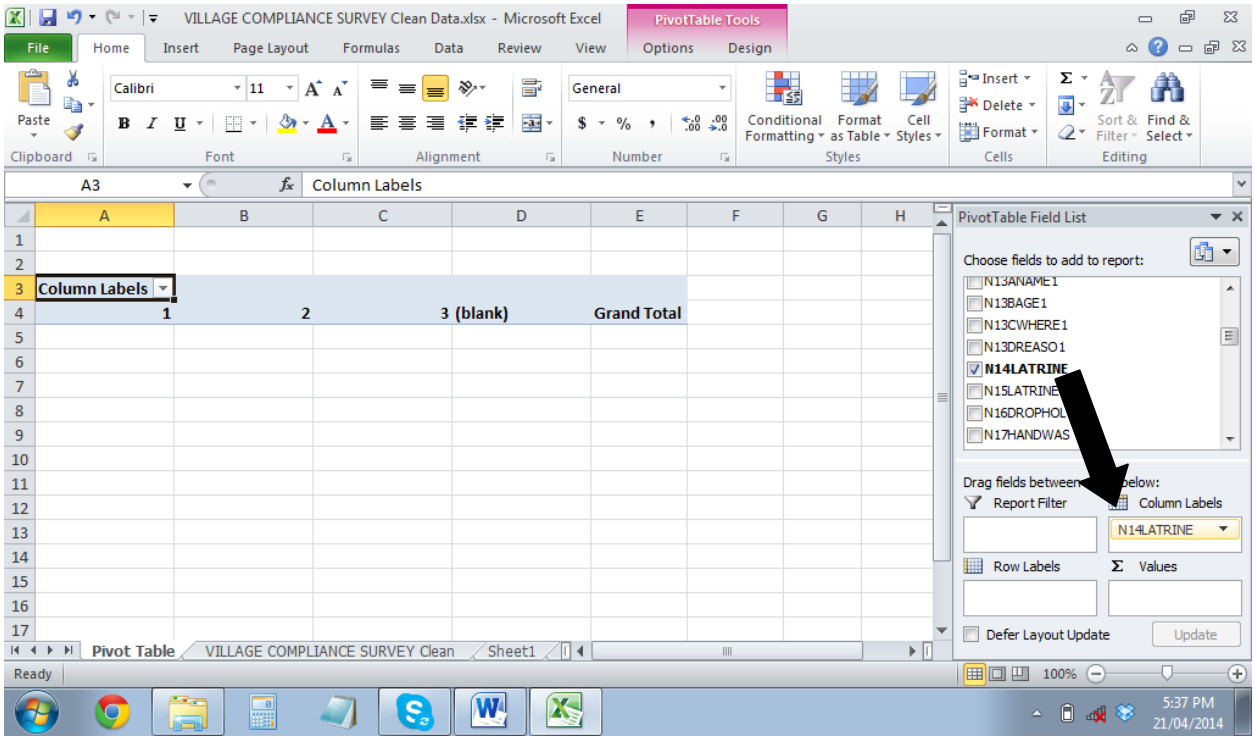
In the following example 969 people answered 1 (Observed Present), 806 people answered 2 (Observed Not Present) and 4 people answered 3 (No permission to observe).



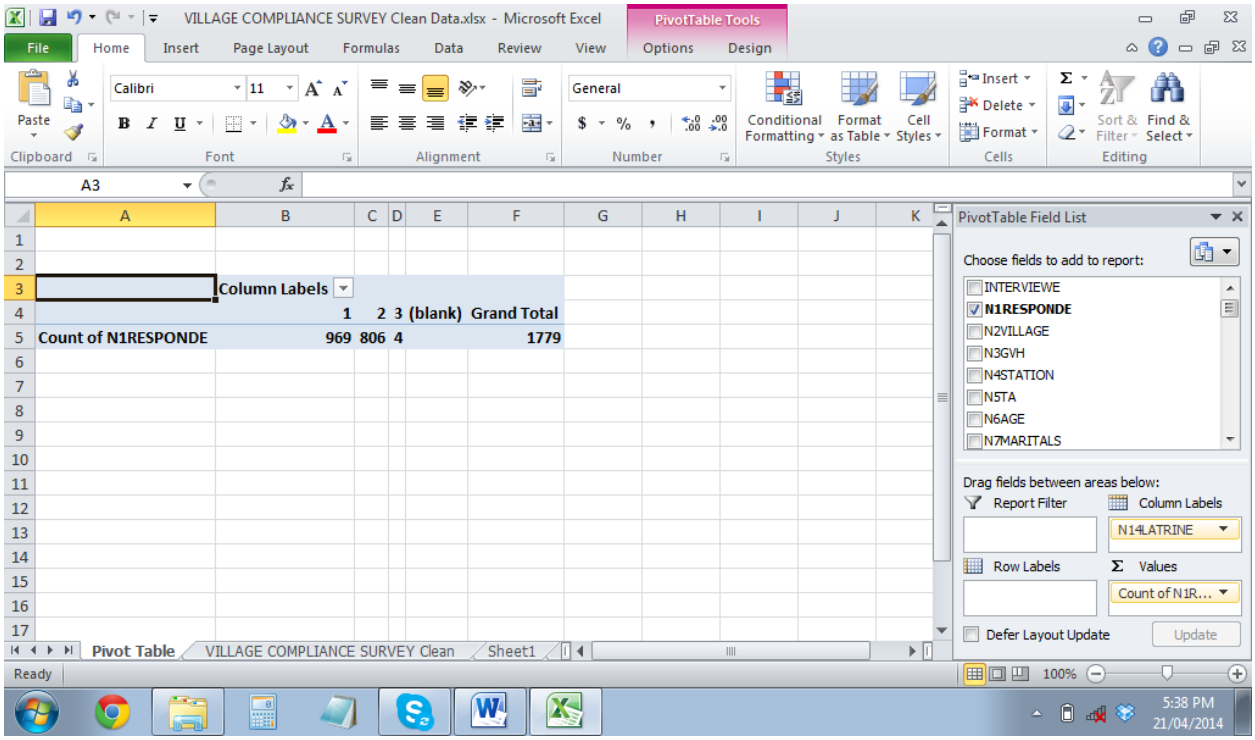
To see the results by station, TA, GVH or Village follow the instructions in section 5.8

5.5.2 Calculating percentages

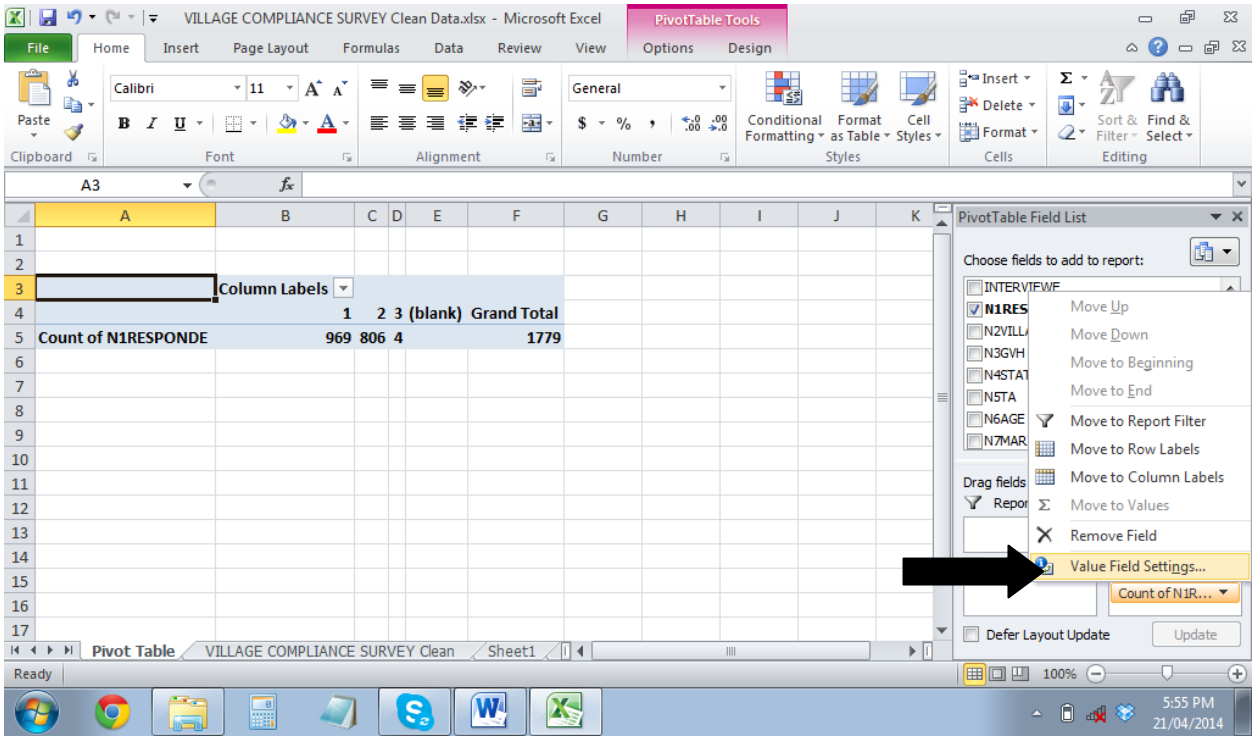
Step 1: Drag the question (in this case Q19. Does the household have a latrine?) into the Column Labels box.



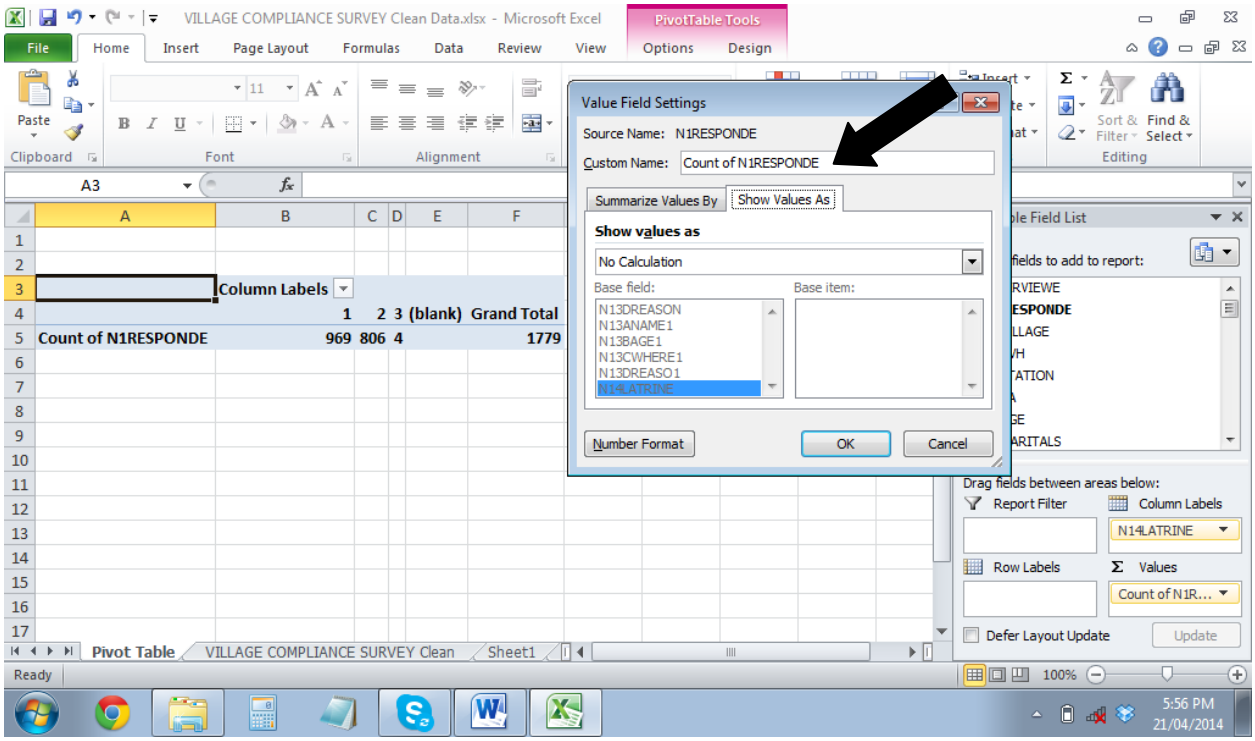
Step 2: Drag the Respondent Code into the Values box.



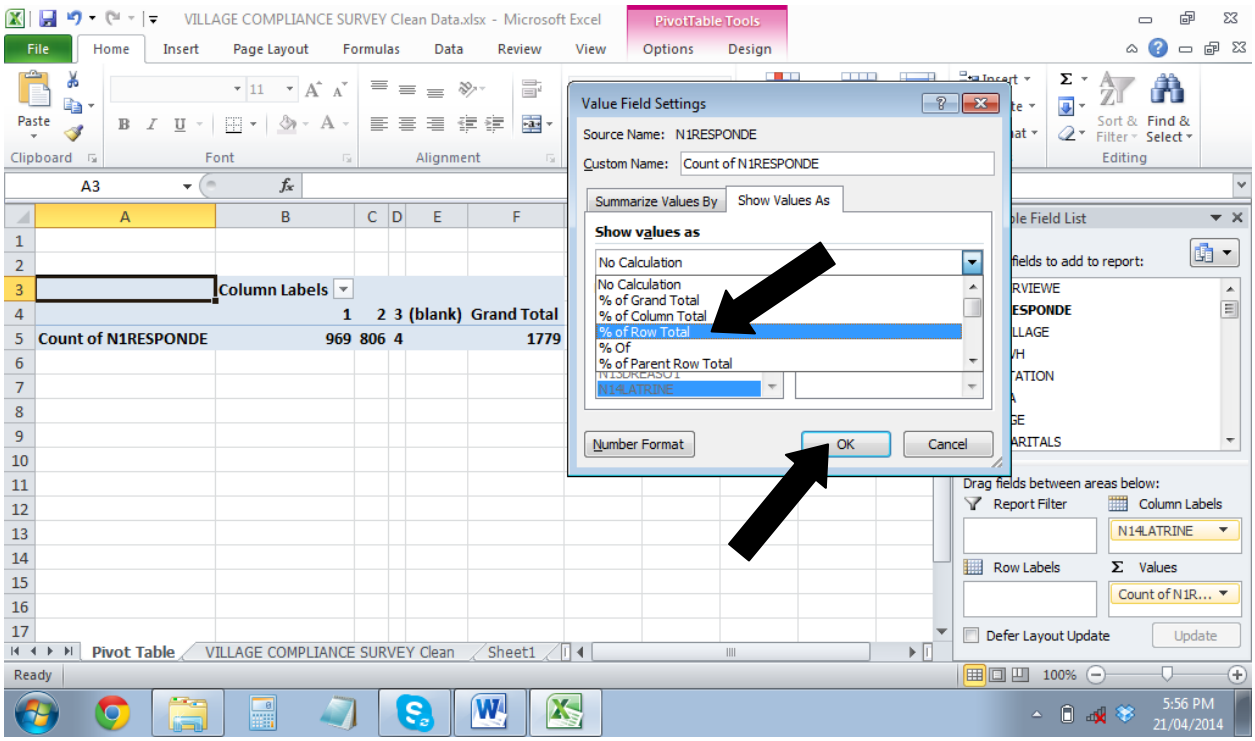
Step 3: Click on the small arrow next to the respondent code in the Values box. Choose Value Field Settings.



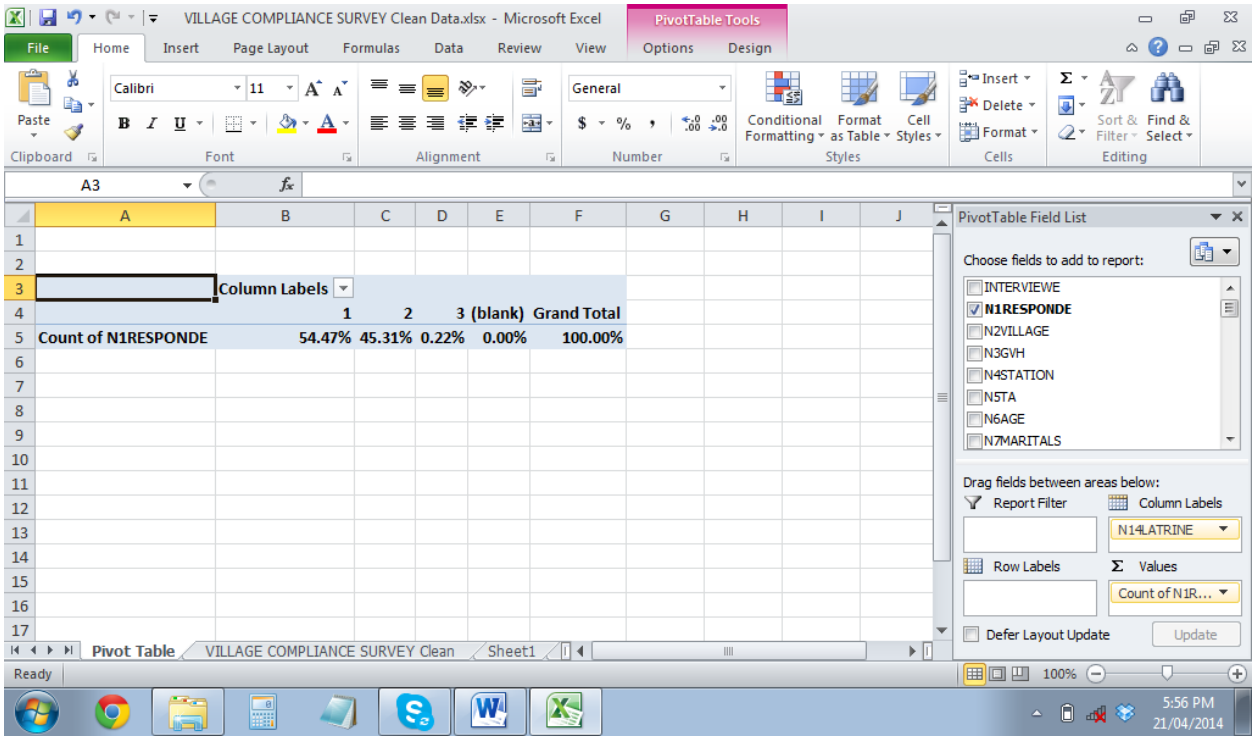
Step 4: Click on the Show Value As tab.



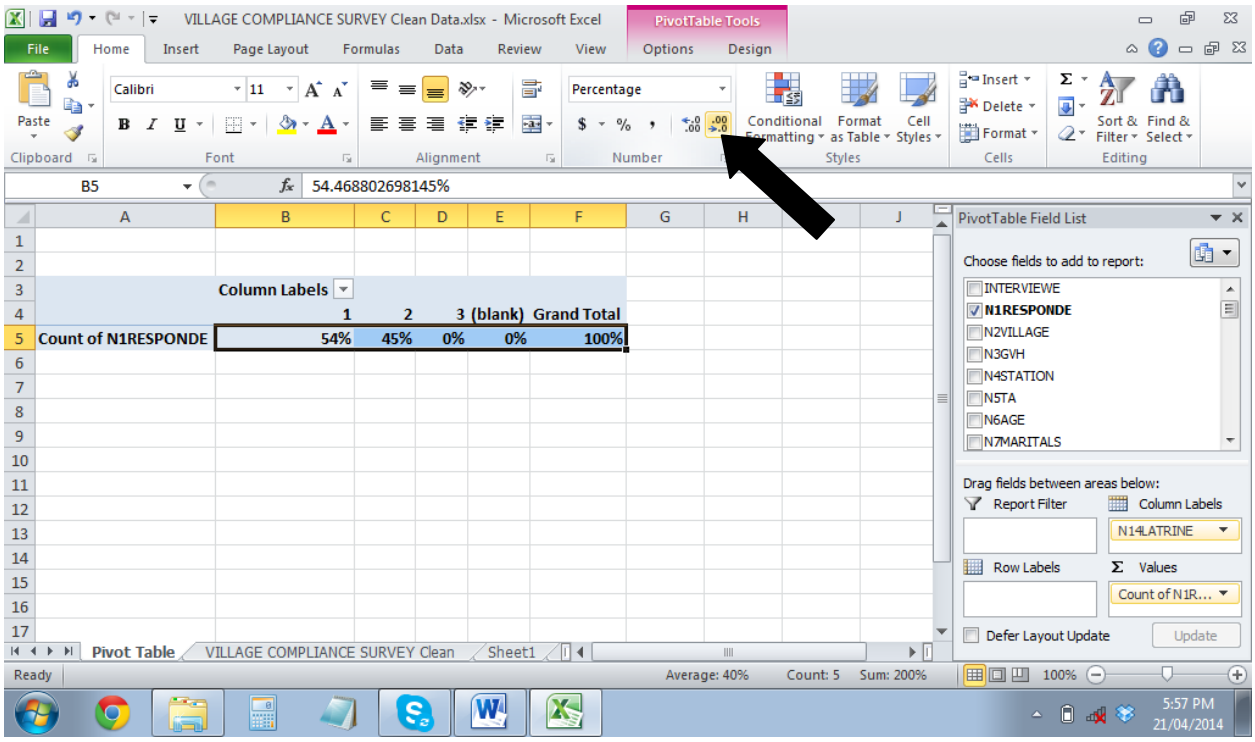
Step 5: Change from No Calculation to % of Row Total. Click OK.



Step 6: The results will now appear as percentages.



Step 7: Reduce the number of decimal points. Do this by selecting the results and clicking on the button shown below. Keep clicking until there are no decimal points.



To see the results by station, TA, GVH or Village follow the instructions in section 5.8

5.6 Questions where the answer is tick all that apply

These questions have a ☐ for the answers. In the spreadsheet there is a separate column for each of the answers. That means you need to do the analysis separately for each answer and then copy all the results into a separate spreadsheet to make one table.

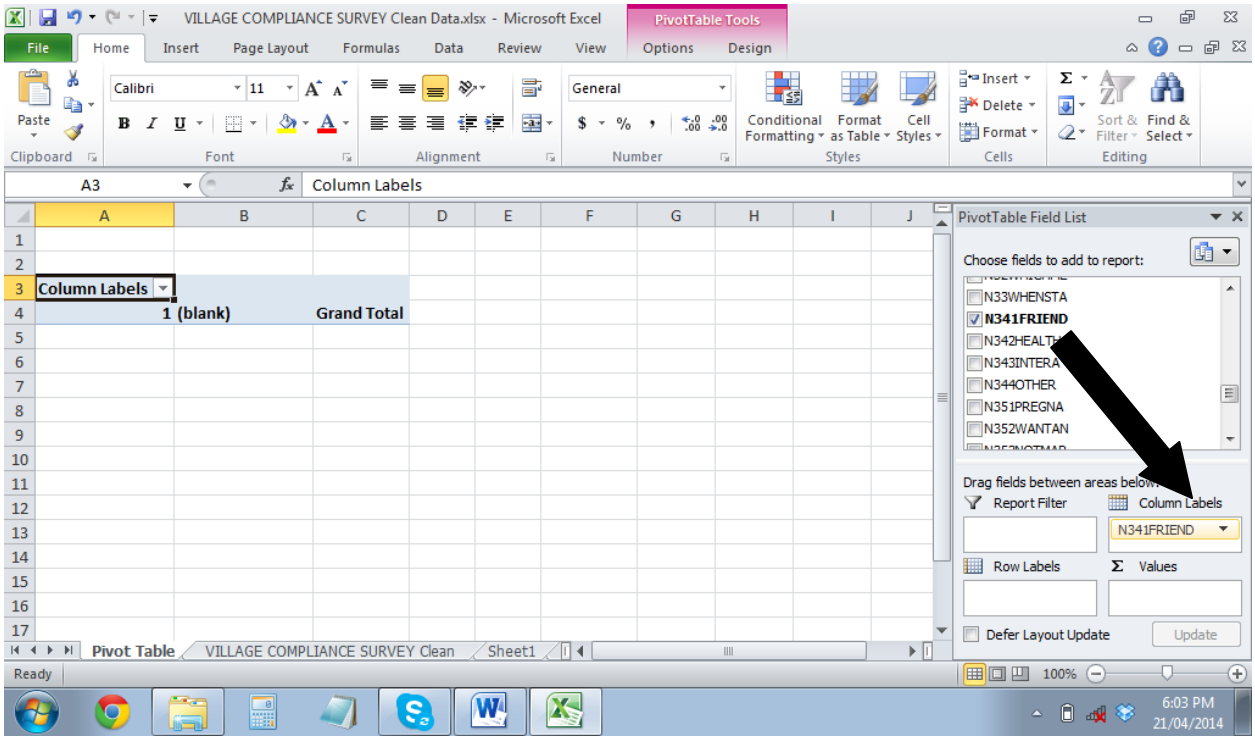
Example

Q34. Who motivated you?

1 ☐ Friend / relative
2 ☐ Health worker (nurse, HSA, etc)
3 ☐ Inter Aide Facilitator
4 ☐ Other

5.6.1 Calculating totals

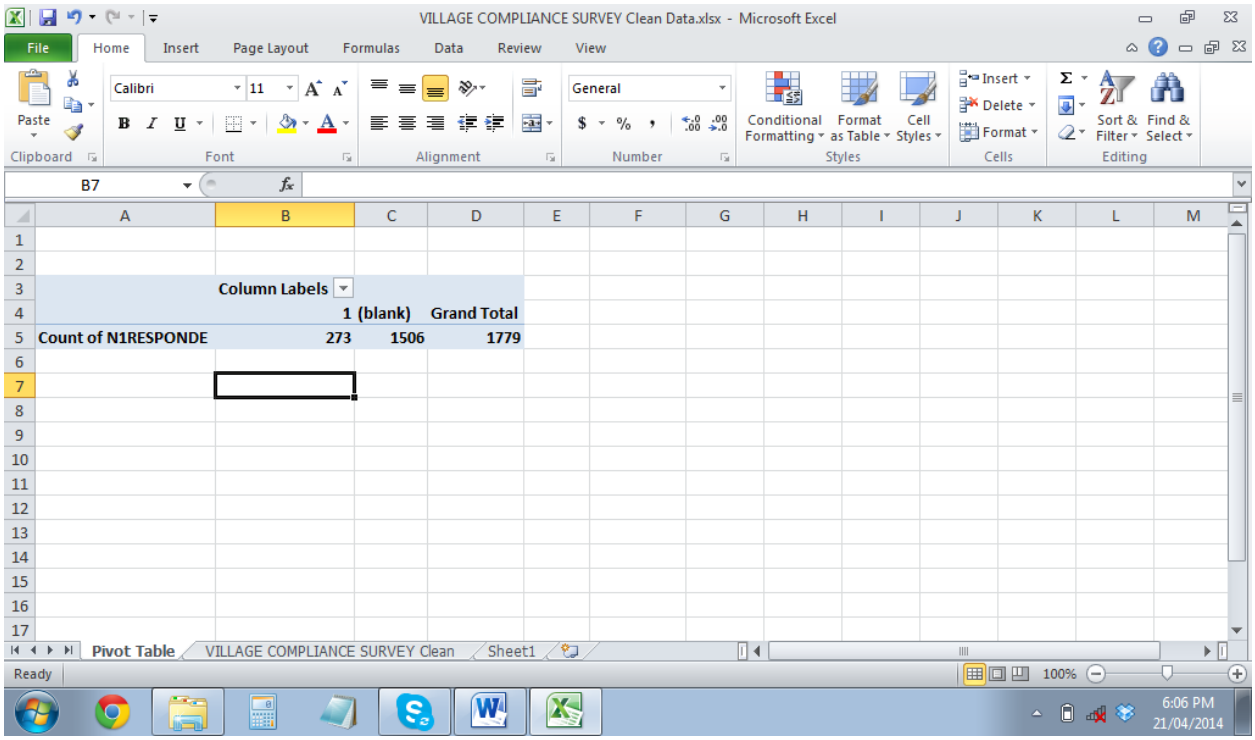
Step 1: Drag the first answer (in this case Q34 answer 1) into the Column Labels box.



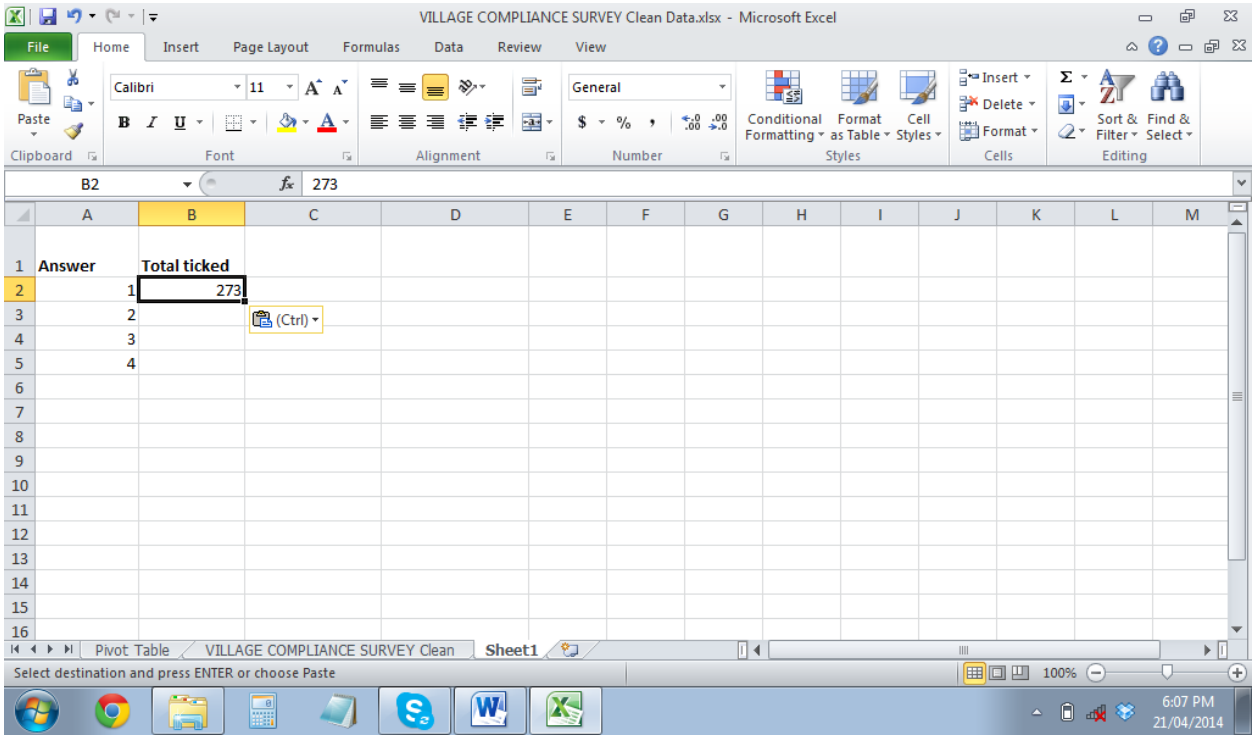
Step 2: Drag the Respondent Code into the Values box. Make sure it says "Count of…". The pivot table will now show the total number of people who ticked that box, and the total number who left it blank.

In the following example 273 people ticked the box for answer 1 (friend / relative) and 1506 people left it blank.

To see the total by station, TA, GVH or Village drag the codes into the Report Filter box or the Row Labels box as described previously.



Step 3: Copy the result for answer 1 into a table on another spreadsheet.

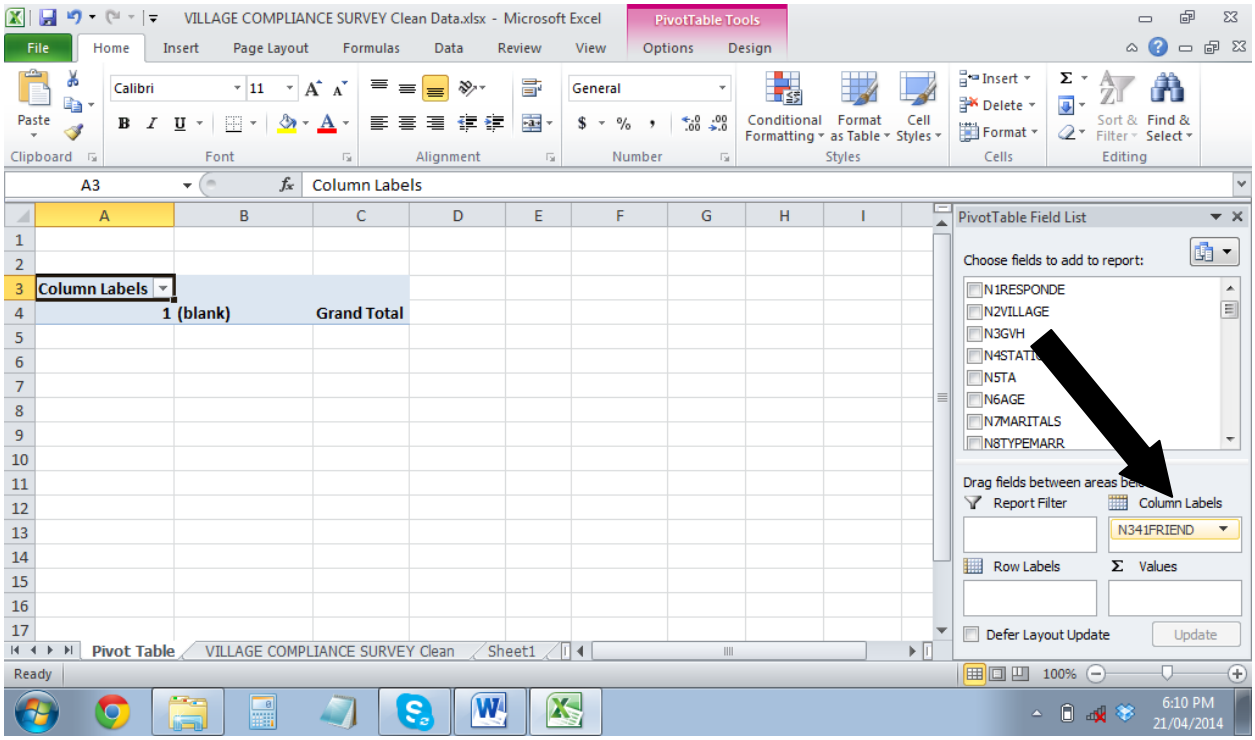


Repeat the same steps for all answers until the table is complete.

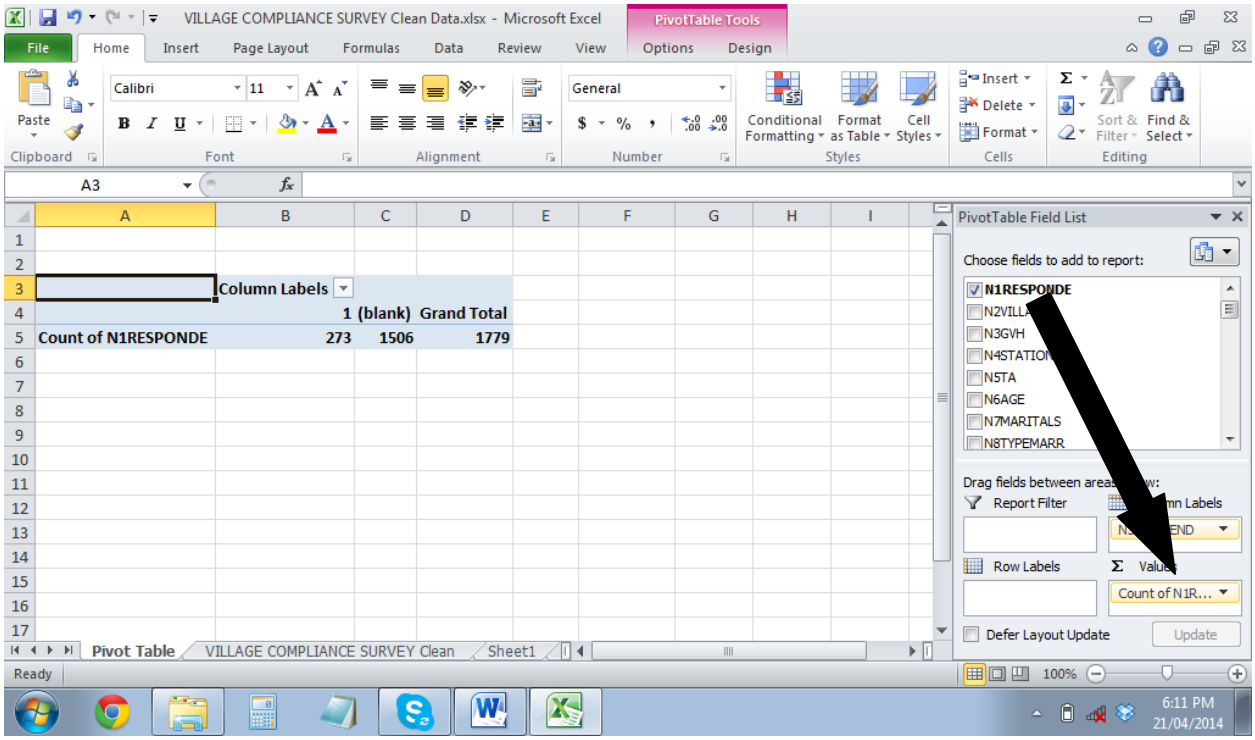
To see the results by station, TA, GVH or Village follow the instructions in section 5.8

5.6.2 Calculating percentages

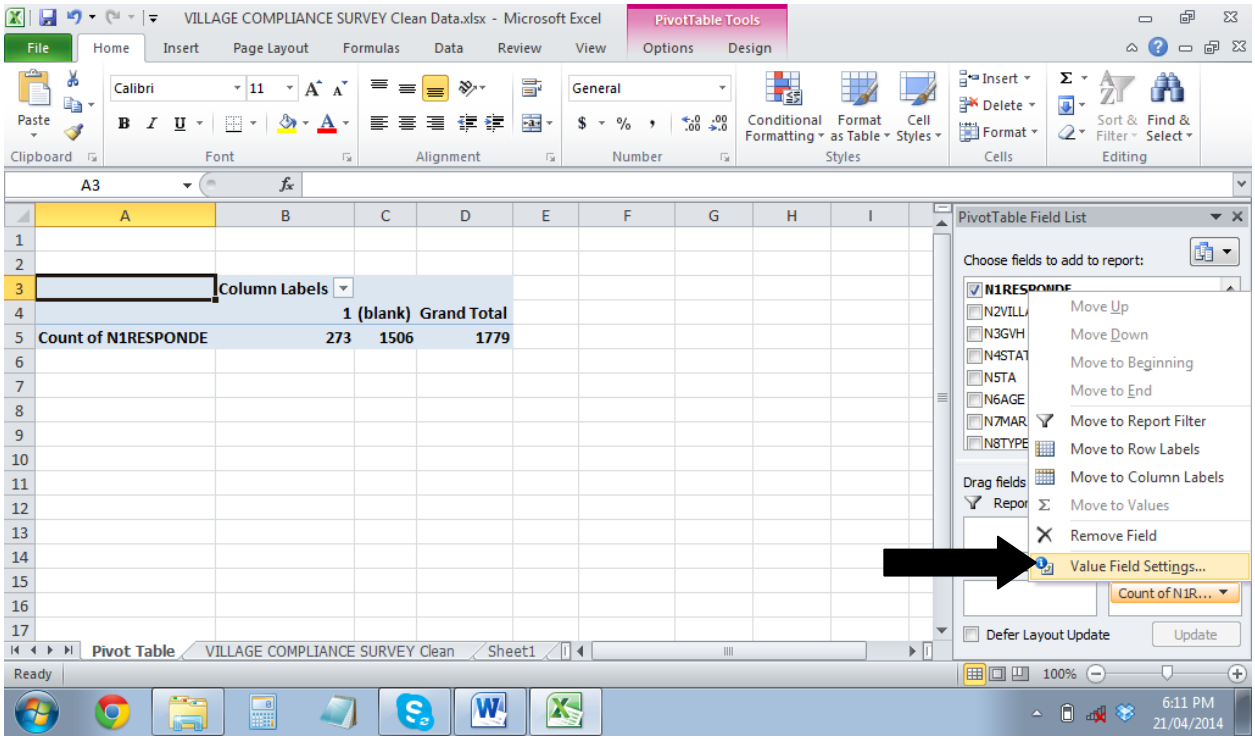
Step 1: Drag the first answer (in this case Q34 answer 1) into the Column Labels box.



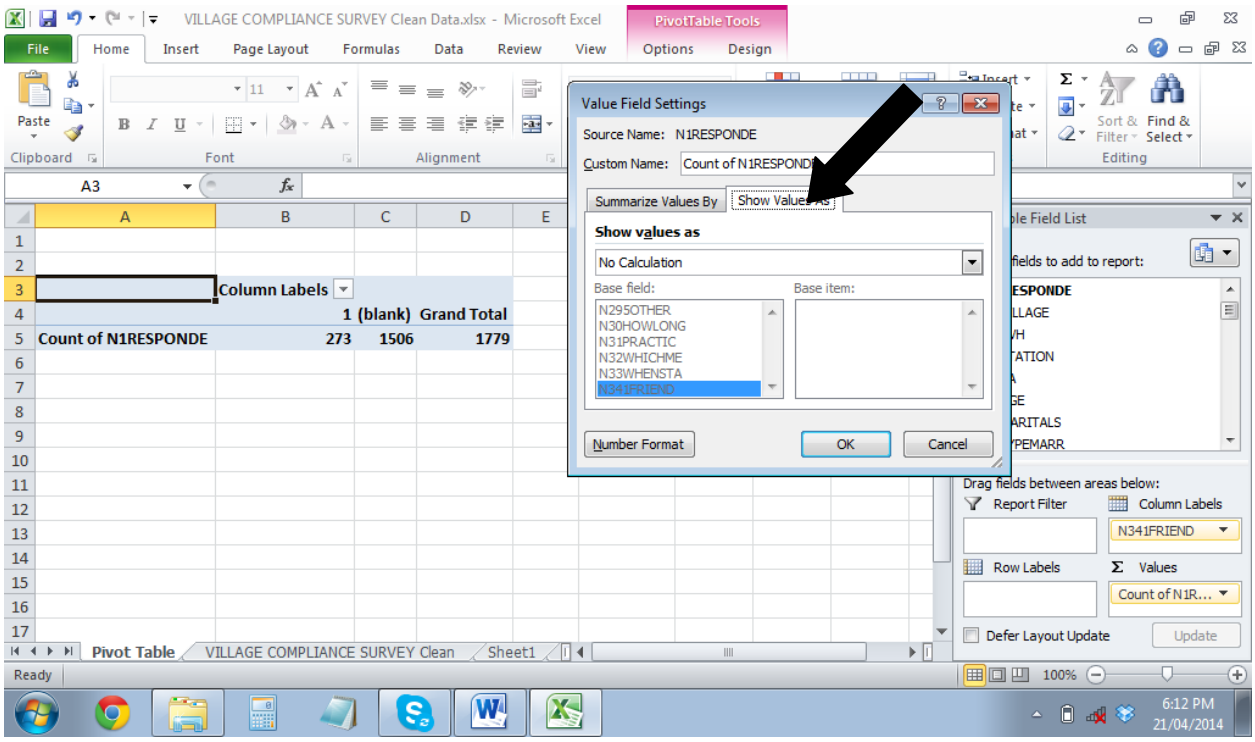
Step 2: Drag the Respondent Code into the Values box. Make sure it says "Count of…".



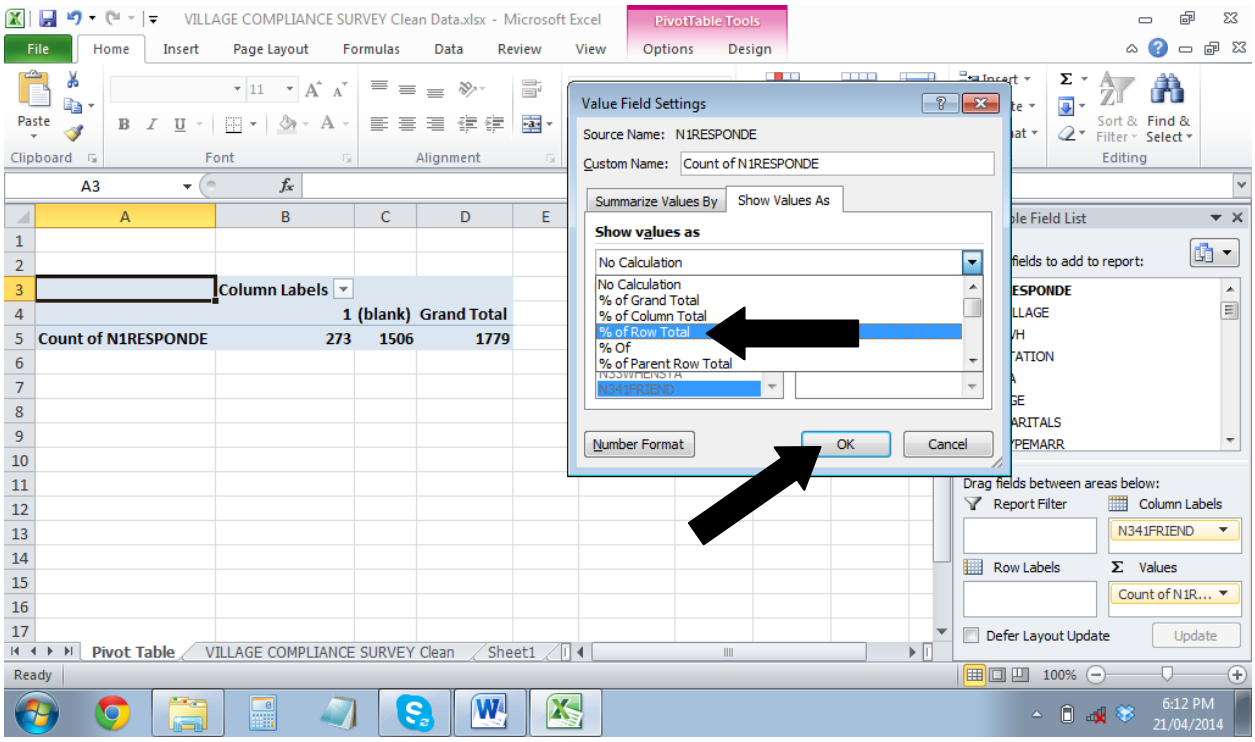
Step 3: Click on the small arrow next to the respondent code in the Values box. Choose Value Field Settings.



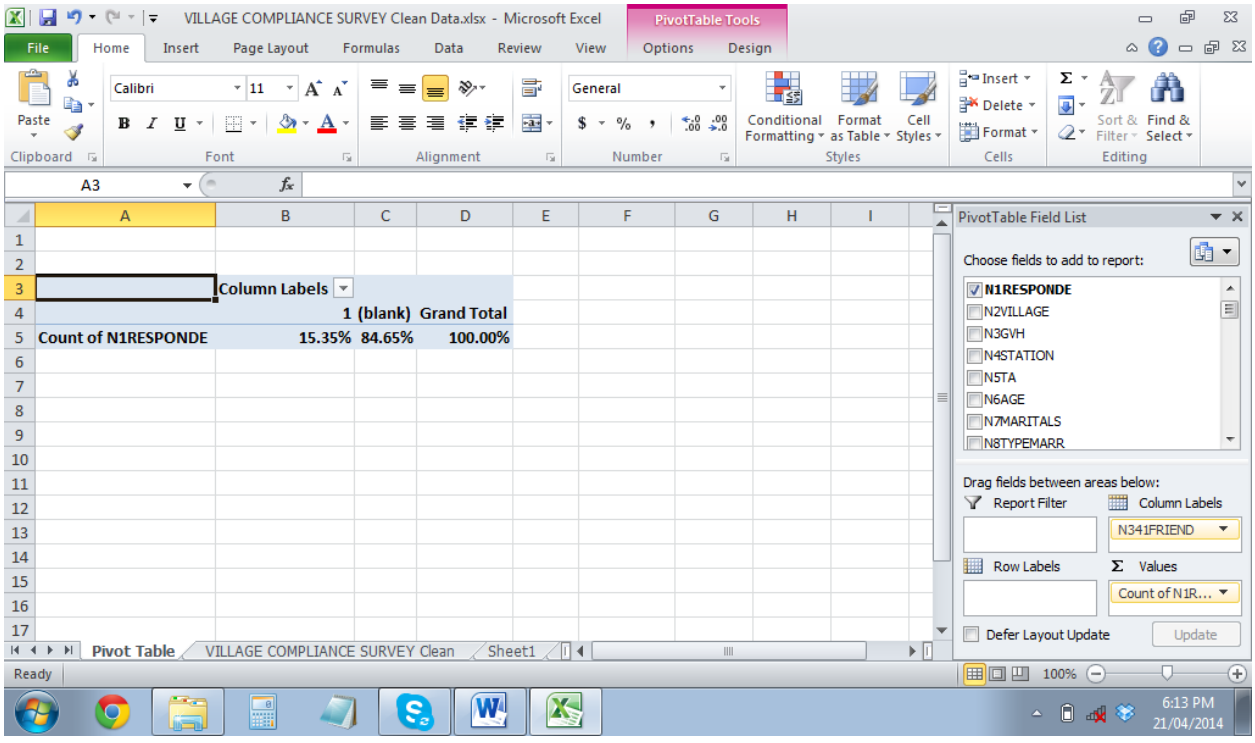
Step 4: Click on the Show Value As tab.



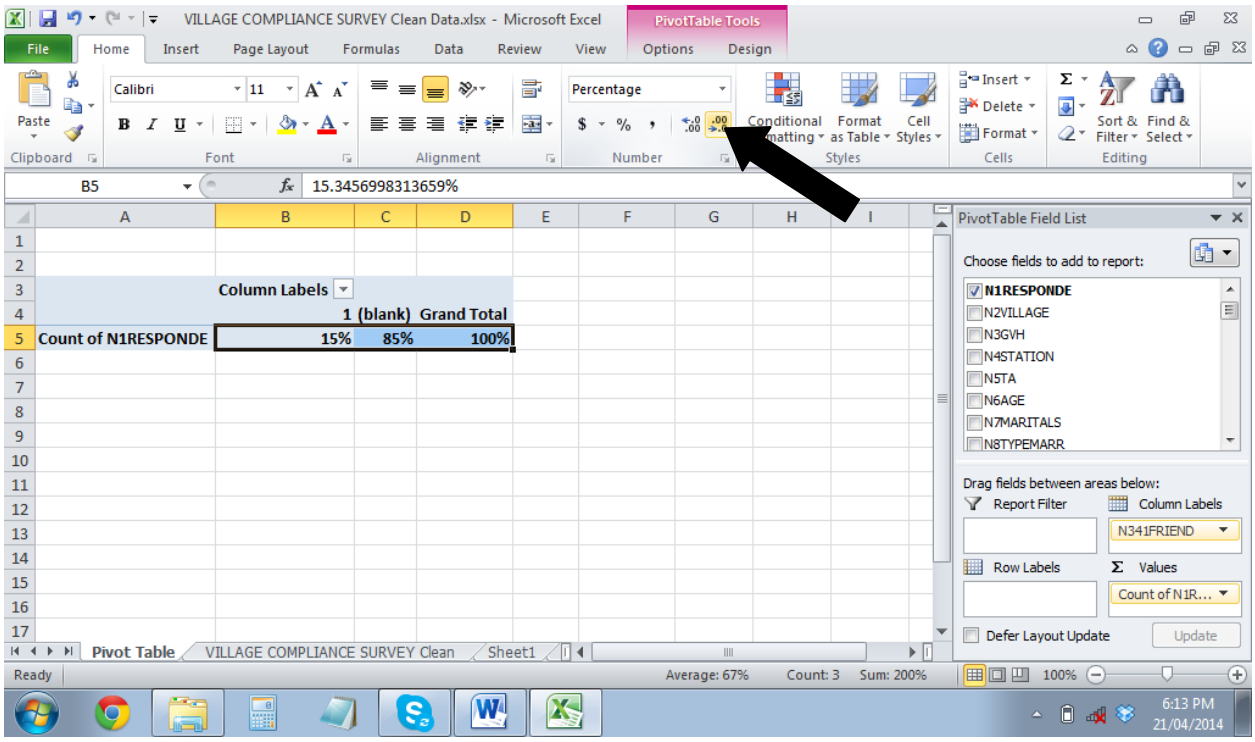
Step 5: Change from No Calculation to % of Row Total. Click OK.



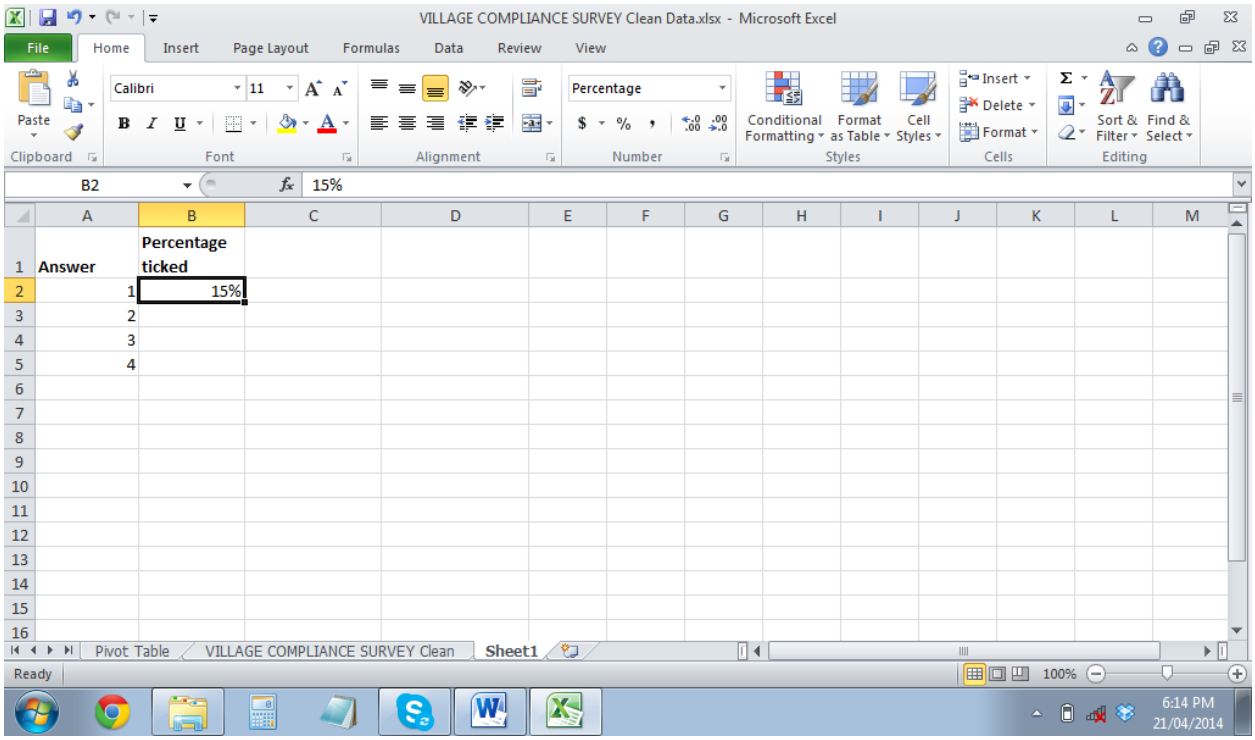
Step 6: The results will now appear as percentages. In the example below 15.35% of people ticked answer 1, 84.65% of people left it blank. *To see the percentages by station, TA, GVH or Village drag the codes into the Report Filter box or the Row Labels box as described previously.*



Step 7: Reduce the number of decimal points. Do this by selecting the results and clicking on the button shown below. Keep clicking until there are no decimal points.



Step 3: Copy the result for answer 1 into a table on another spreadsheet.



Repeat the same steps for all answers until the table is complete.

To see the results by station, TA, GVH or Village follow the instructions in section 5.8

5.7 Questions with jumps

A jump is indicated by a ➔. When a question is jumped the jump needs to be considered in the analysis.

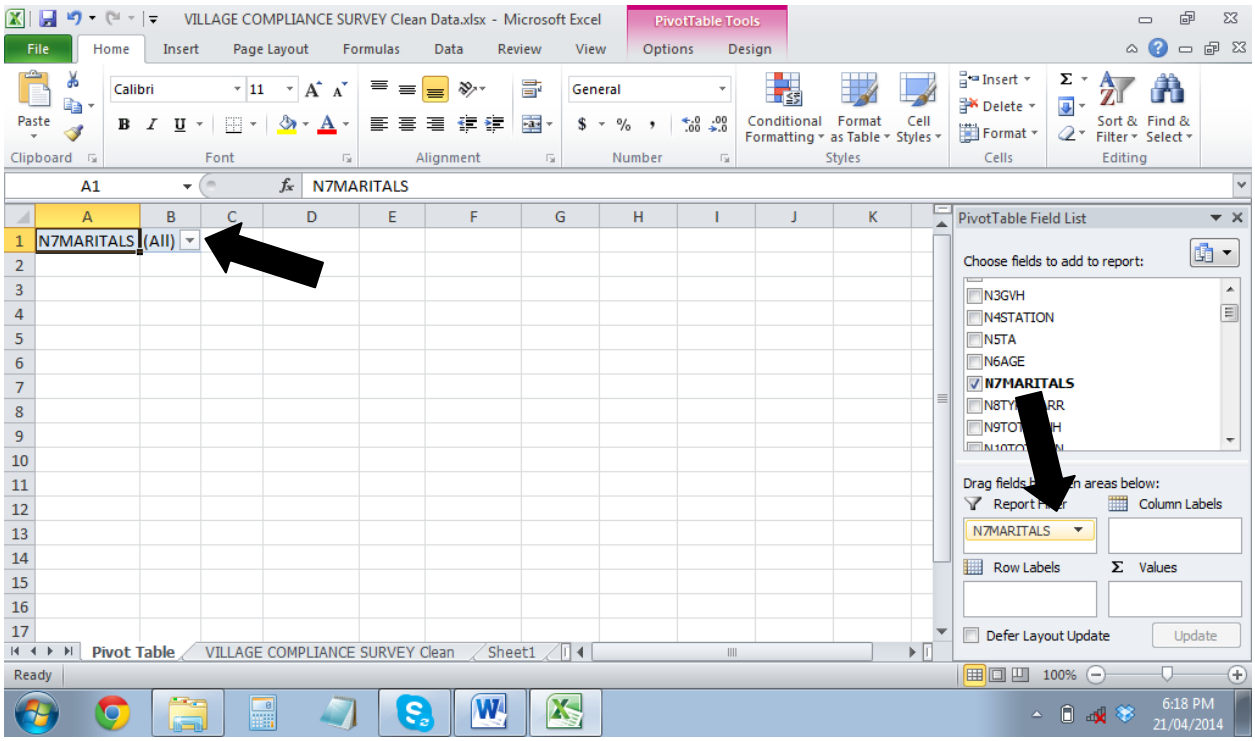
Example

Q7. Marital status
1 ○ Married
2 ○ Single ➔ **Q9**
3 ○ Divorced ➔ **Q9**
4 ○ Widow ➔ **Q9**
5 ○ Other ➔ **Q9**

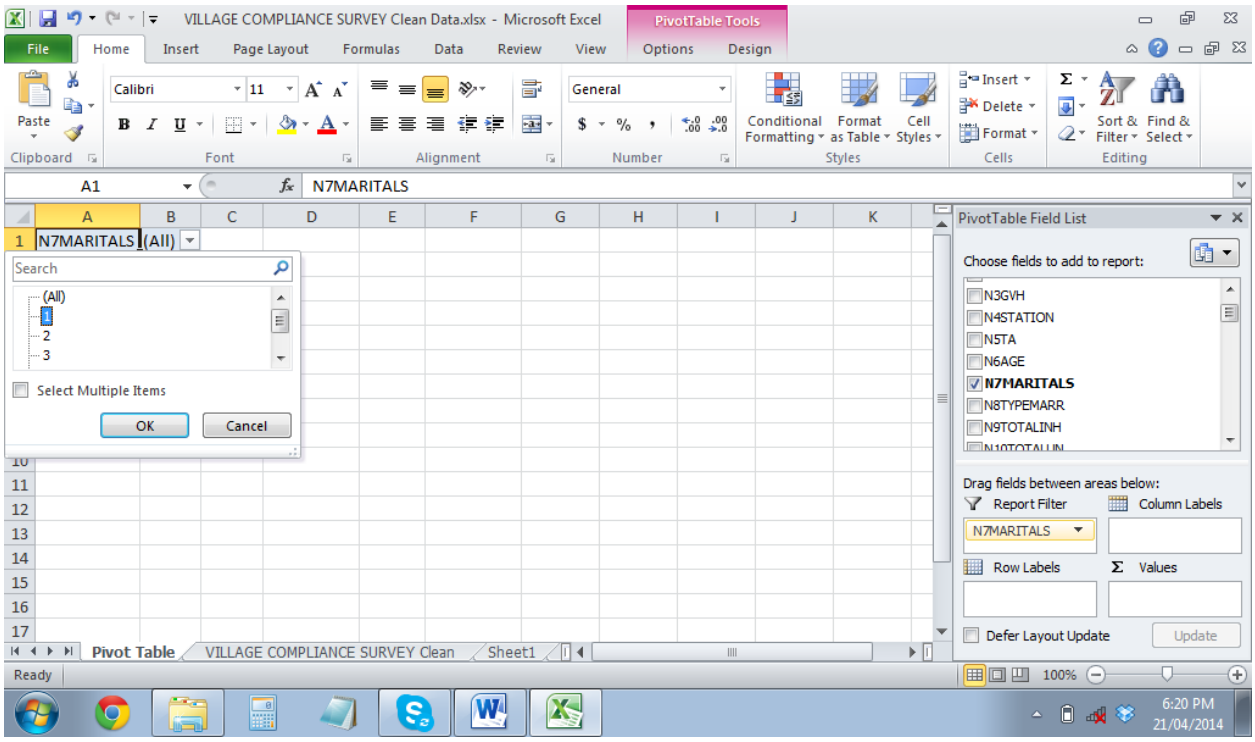
Q8. Type of marriage
1 ○ Monogamous
2 ○ Polygamous

In this example the results for Q8 should only be calculated for households where the answer for Q7 is 1 (Married). This can be achieved by using a filter.

Step 1: Drag the question with the jump (in this case Q7) into the Report Filter box. It will appear as a filter at the top of the Pivot Table.



Step 2: Click the arrow button on the filter. Select the answer without the jump (in this case answer 1). Click OK.

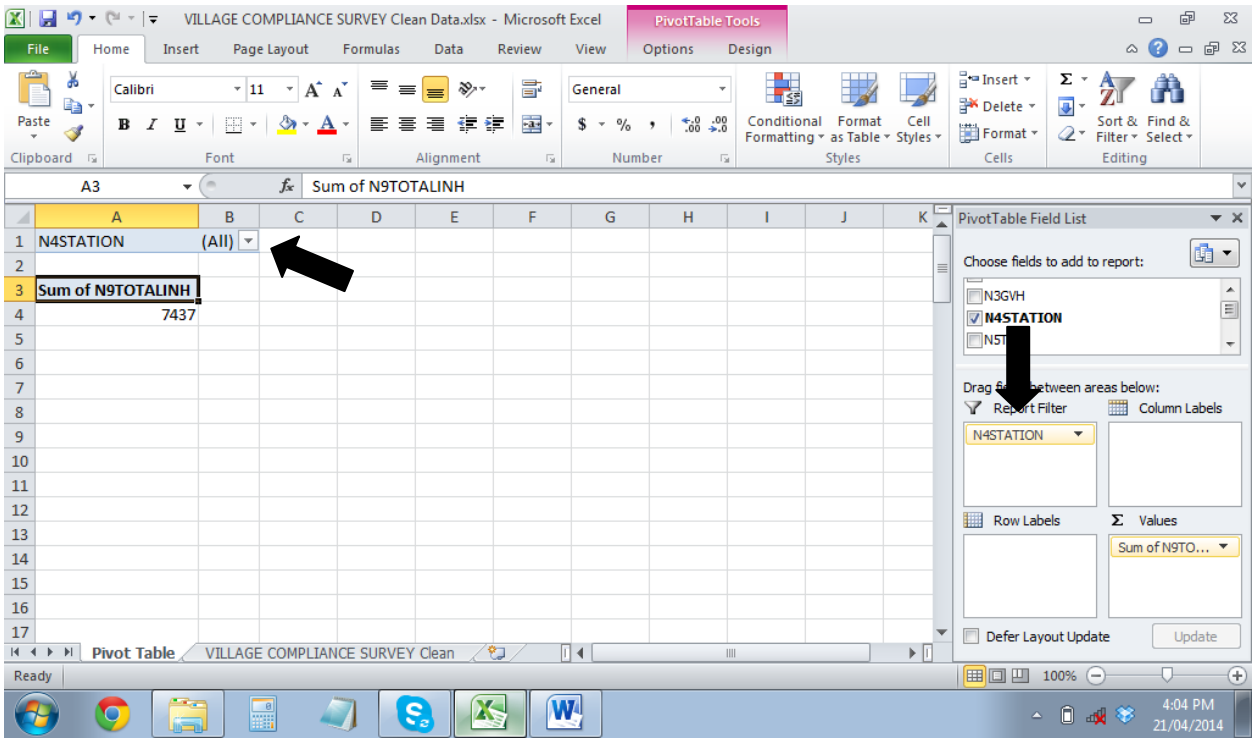


Step 3: Proceed with your analysis of the question (in this case Q8) as usual. Only households that answered 1 to Q7 will be included in the analysis. Remember to remove the filter after you finish analysing Q8, before moving on to the next question.

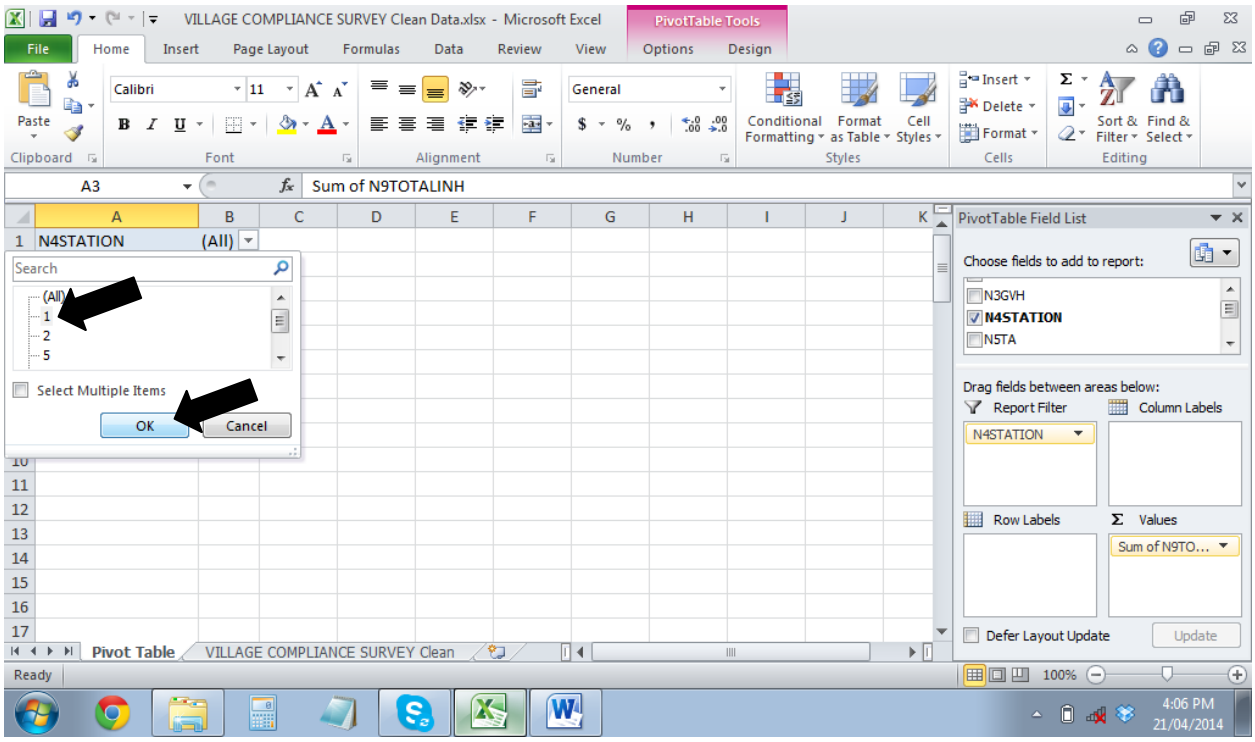
5.8 Showing results by Station, TA, GVH, or Village

Complete your analysis using the previous steps. Then follow these steps to break the results down by Station, TA, GVH and Village.

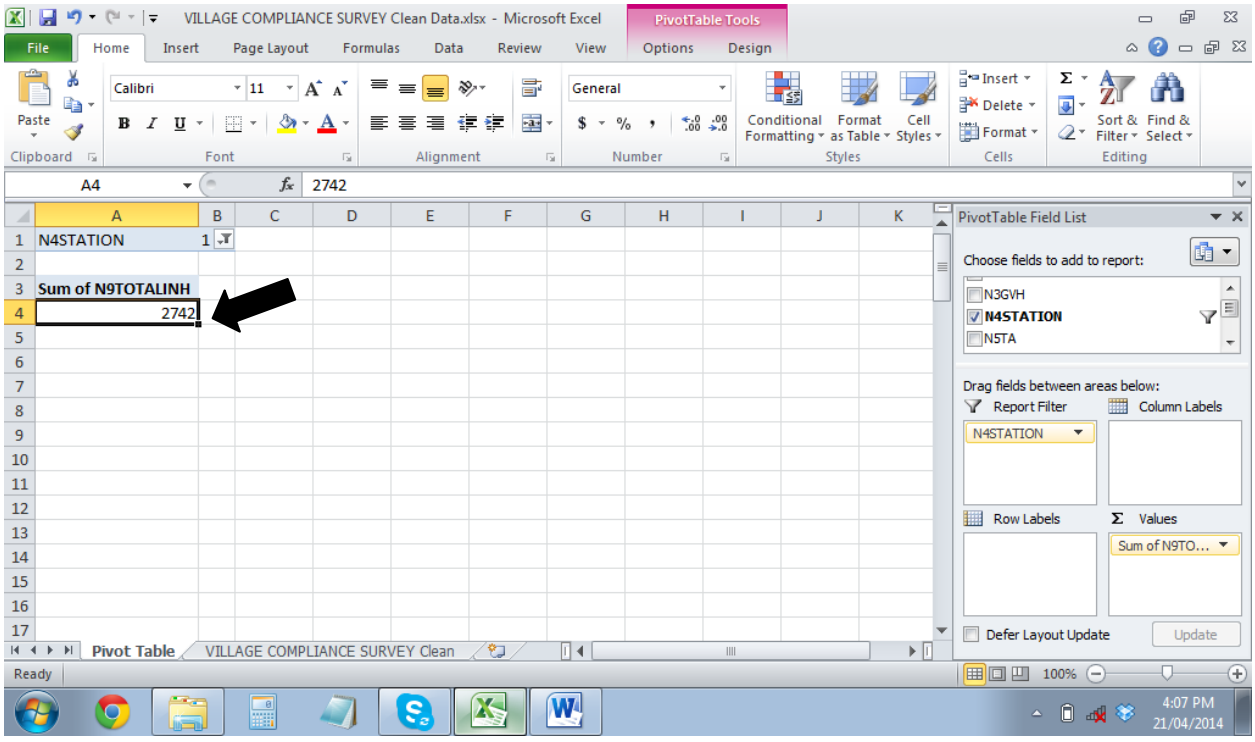
Step 1: To see the results for a particular station, drag the station code into the Report Filter box. The filter will appear at the top of the pivot table.



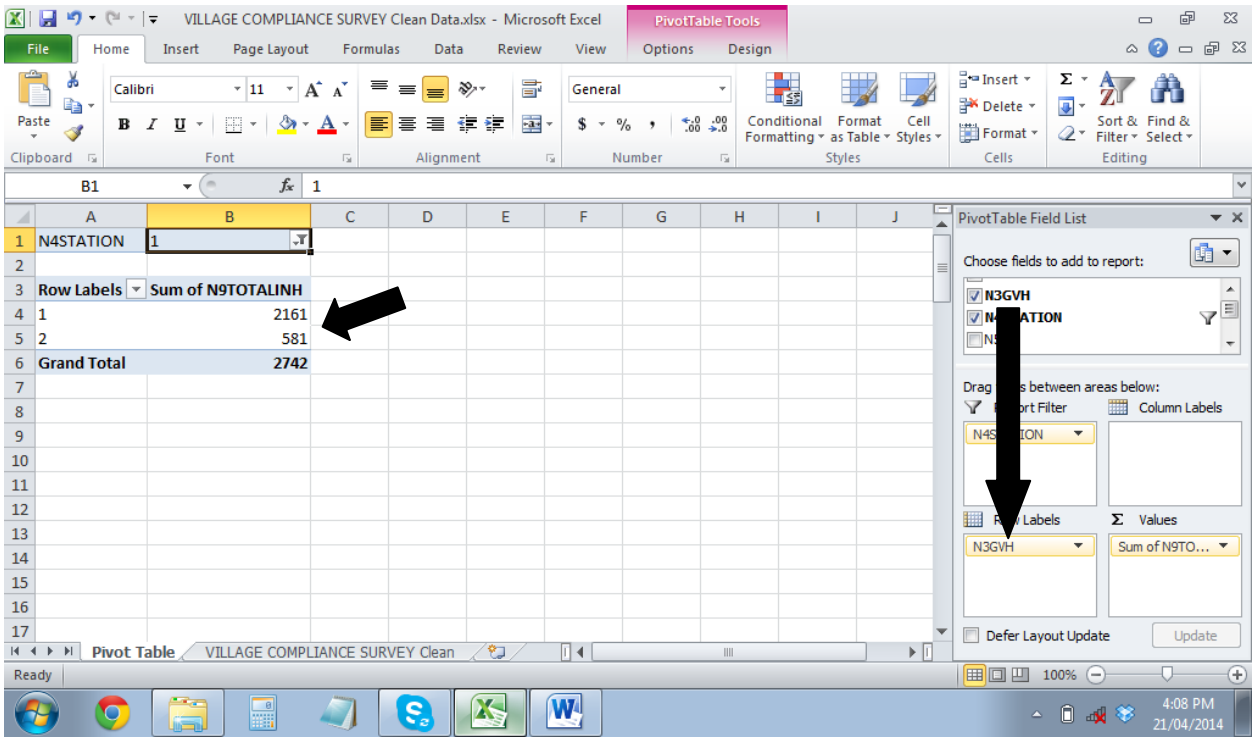
Step 6: Click on the arrow button next to the filter. Choose the code for the station you would like to see results for (in this case I have chosen station #1, Mlodzenzi). Click OK.



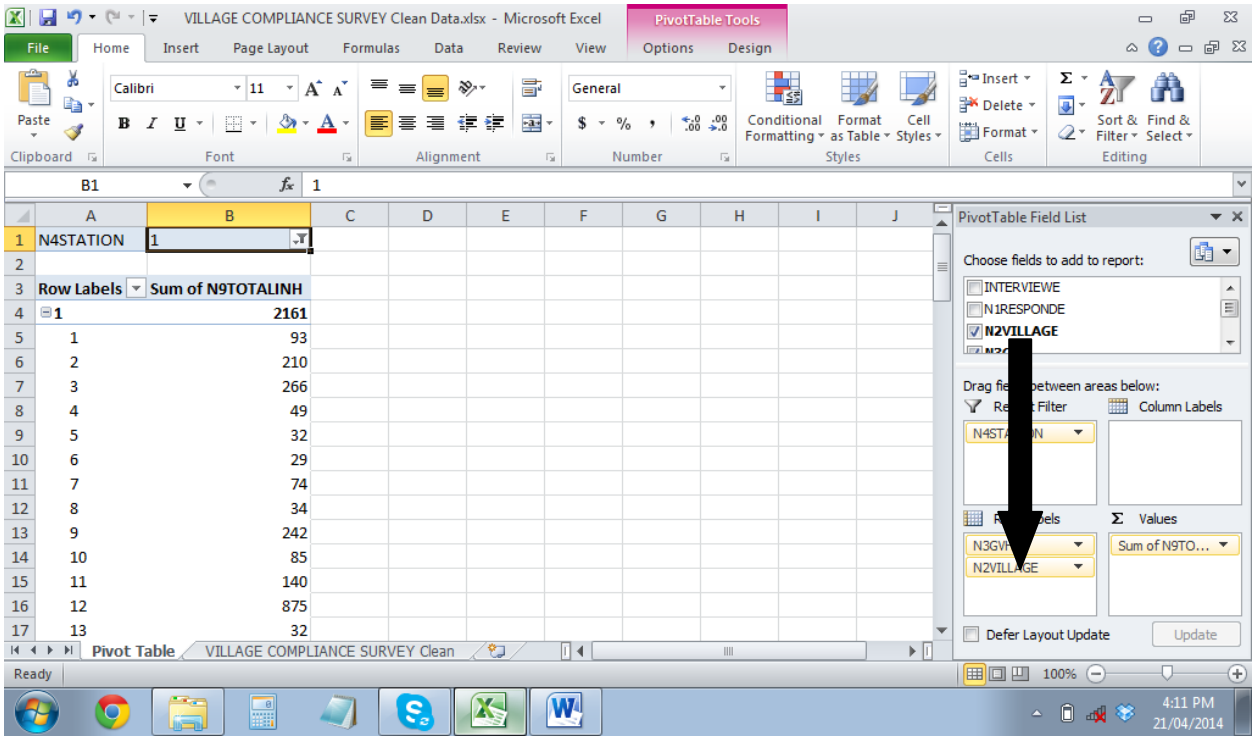
Step 6: The pivot table has now updated to show the result for all households in station #1.



Step 7: To see the results by GVH, drag the GVH code into the Row Labels box. The pivot table will update to show the GVH codes in the first column and the results in the second column. In this example there are two GVHs, GVH Chizumba (#1 with total inhabitants of 2161) and GVH Mkakaula (#2 with total inhabitants of 581).



Step 8: To see the results by village under each GVH, drag the village code into the Row Labels box so that it is under the GVH code. The table will update to show results for each village code under the GVH. For example, under GVH #1 (Chizumba) there is village #1 (Mtibu) with a total inhabitants of 93.



6 Mapping

6.1 Collecting GPS coordinates for locations

The first step in creating a map is to get GPS coordinates for all the locations that you want to map (e.g. villages, water points, schools, etc). The easiest way to get GPS coordinates is using a smart phone.

6.1.1 Download a GPS app

First download a GPS app to the phone. If you have an Android phone then a good free app is **GPS-Simple for Android**. You can download it here:
https://play.google.com/store/apps/details?id=com.andreabaccega.simplegps

The app must product decimal GPS coordinates, not coordinates in degrees, minutes and seconds.

- **The coordinates should look like this: 33.1675623, -42.7462729**
- **They should NOT look like this: 40° 26′ 46″ N 79° 58′ 56″ W**

If your app only produces GPS coordinates in degrees then you will need to use a converter like this one to turn them into decimals: http://transition.fcc.gov/mb/audio/bickel/DDDMMSS-decimal.html

6.1.2 Go to the locations and record the GPS coordinates

Once you've got your phone you should go to each location you want to map and use the GPS app to get the latitude and longitude. If the location is very remote you might need to move around a bit until your phone can get a lock on it.

Write the latitude and longitude down on a record sheet, along with the name of the location and what it is (village, health centre, school, borehole, etc). An example of a record sheet is shown below:

Latitude and longitude come from the GPS app on your phone

#	Location Name	Type	Latitude	Longitude
1	Mgoola	Village	-14.18748	33.75930
2	Chingoyo	Village	-14.18746	33.75925
3	Zemba	Village	-14.18753	33.75953
4	Chindza	Primary school	-14.18748	33.75999
5	Mkenda	Secondary school	-14.18743	33.75991
6				
7				
8				

Include all the locations that are relevant for your project. (villages, schools, health centres, boreholes, government offices, etc)

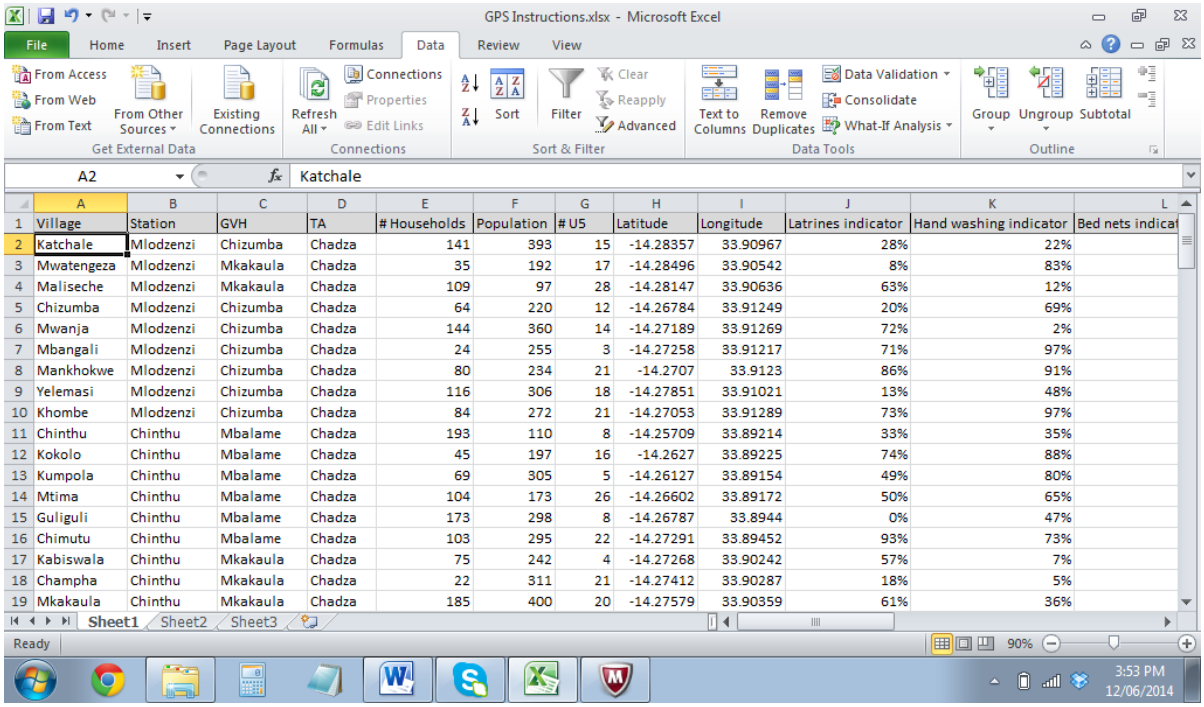
6.2 Preparing the spreadsheet

Before you can make a map you need to prepare your data in a spreadsheet. The spreadsheet should have one heading row, and then one row for each village. For the child health project the columns should be:

- Village
- Station
- GVH
- TA
- # Households
- Population
- # U5
- Latitude
- Longitude
- Latrines indicator (%)
- Hand washing indicator (%)
- Bed nets indicator (%)
- Health seeking indicator (%)
- Family planning indicator (%)
- Safe delivery indicator (%)
- Village clinics indicator (%)

Make sure all station, GVH and TA names are spelt the same way every time. DO NOT merge cells or leave any blank rows or columns. DO NOT use formulas. If data is missing for a village then leave the cell blank, do not write any text like "n/a" or "unknown".

When you are finished the spreadsheet should look like this:

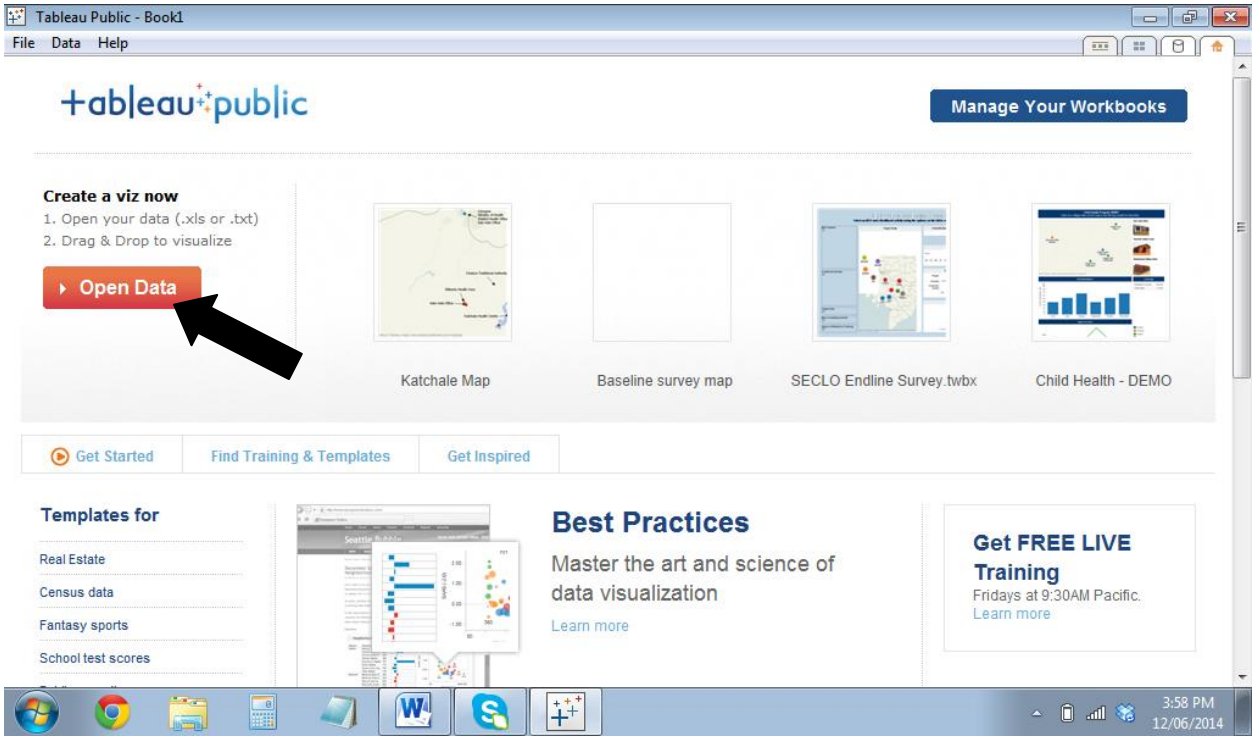


	Village	Station	GVH	TA	# Households	Population	# U5	Latitude	Longitude	Latrines indicator	Hand washing indicator	Bed nets indicat
2	Katchale	Mlodzenzi	Chizumba	Chadza	141	393	15	-14.28357	33.90967	28%	22%	
3	Mwatengeza	Mlodzenzi	Mkakaula	Chadza	35	192	17	-14.28496	33.90542	8%	83%	
4	Maliseche	Mlodzenzi	Mkakaula	Chadza	109	97	28	-14.28147	33.90636	63%	12%	
5	Chizumba	Mlodzenzi	Chizumba	Chadza	64	220	12	-14.26784	33.91249	20%	69%	
6	Mwanja	Mlodzenzi	Chizumba	Chadza	144	360	14	-14.27189	33.91269	72%	2%	
7	Mbangali	Mlodzenzi	Chizumba	Chadza	24	255	3	-14.27258	33.91217	71%	97%	
8	Mankhokwe	Mlodzenzi	Chizumba	Chadza	80	234	21	-14.2707	33.9123	86%	91%	
9	Yelemasi	Mlodzenzi	Chizumba	Chadza	116	306	18	-14.27851	33.91021	13%	48%	
10	Khombe	Mlodzenzi	Chizumba	Chadza	84	272	21	-14.27053	33.91289	73%	97%	
11	Chinthu	Chinthu	Mbalame	Chadza	193	110	8	-14.25709	33.89214	33%	35%	
12	Kokolo	Chinthu	Mbalame	Chadza	45	197	16	-14.2627	33.89225	74%	88%	
13	Kumpola	Chinthu	Mbalame	Chadza	69	305	5	-14.26127	33.89154	49%	80%	
14	Mtima	Chinthu	Mbalame	Chadza	104	173	26	-14.26602	33.89172	50%	65%	
15	Guliguli	Chinthu	Mbalame	Chadza	173	298	8	-14.26787	33.8944	0%	47%	
16	Chimutu	Chinthu	Mbalame	Chadza	103	295	22	-14.27291	33.89452	93%	73%	
17	Kabiswala	Chinthu	Mkakaula	Chadza	75	242	4	-14.27268	33.90242	57%	7%	
18	Champha	Chinthu	Mkakaula	Chadza	22	311	21	-14.27412	33.90287	18%	5%	
19	Mkakaula	Chinthu	Mkakaula	Chadza	185	400	20	-14.27579	33.90359	61%	36%	

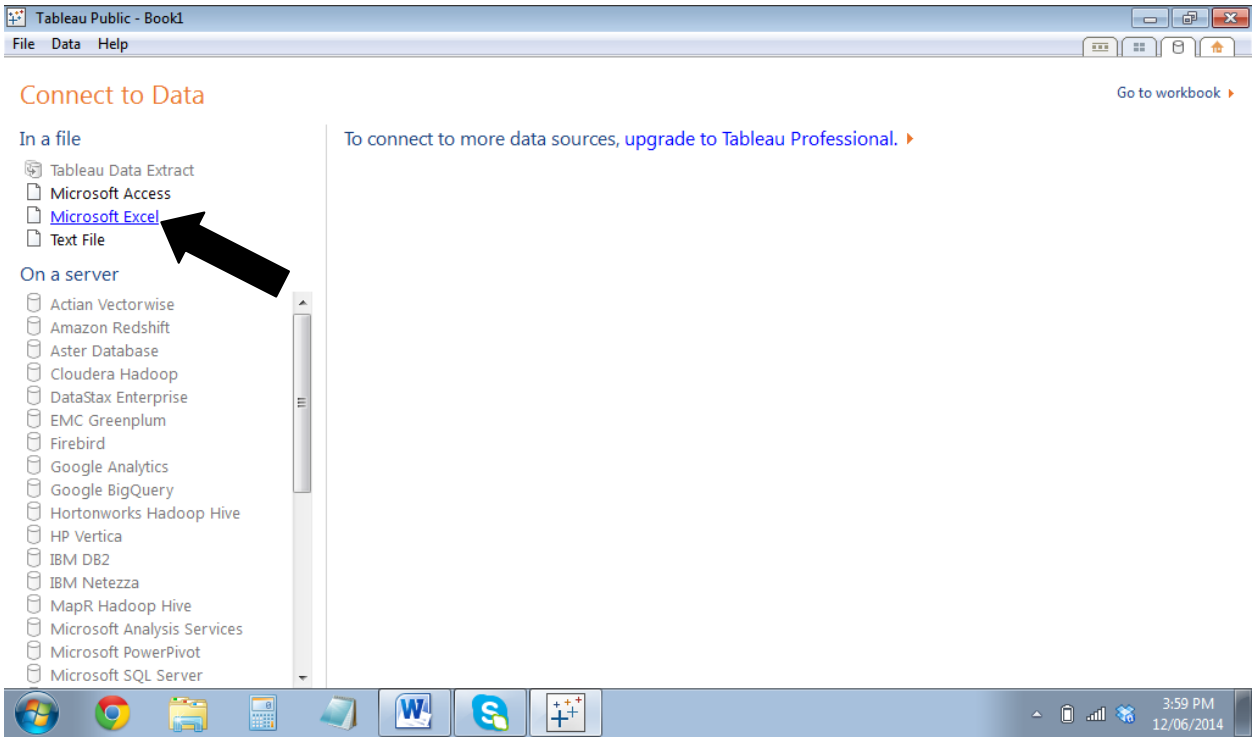
6.3 Making a basic map

Step 1: Connect to the internet. To make a map you MUST be connected to the internet for all steps (the faster the better).

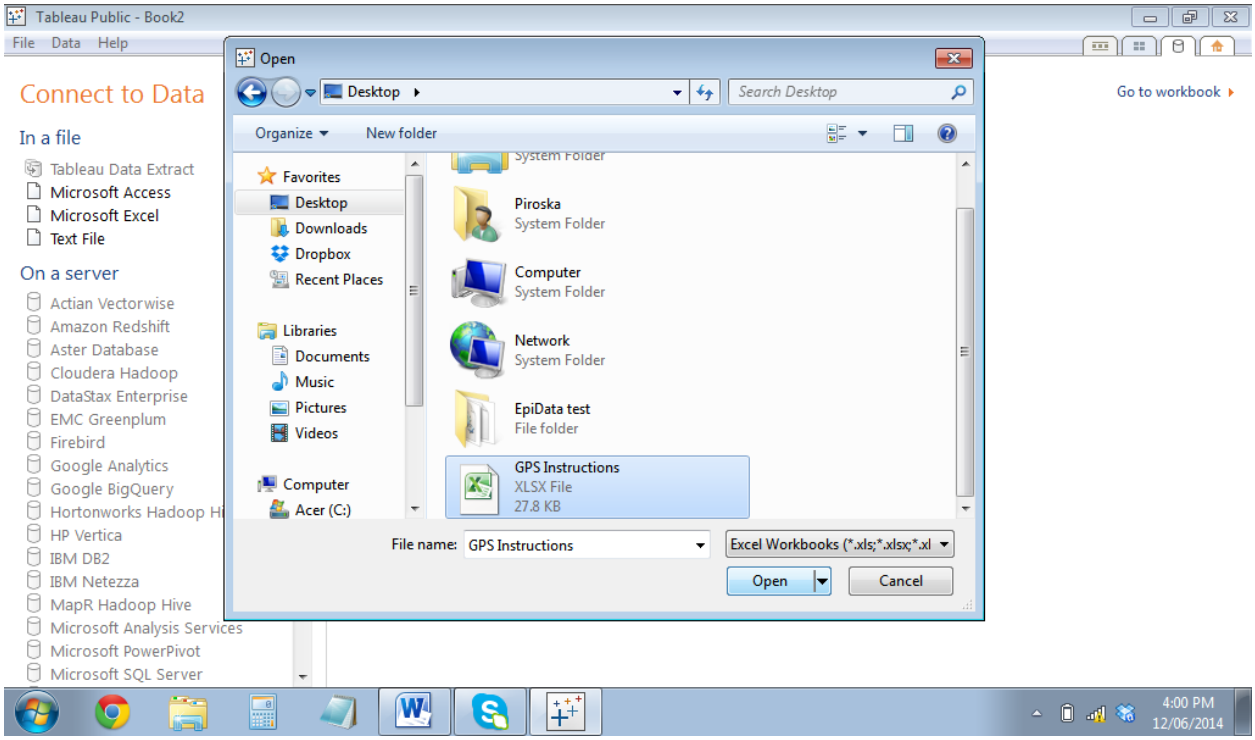
Step 2: Open Tableau Public. Click on the Open Data button.



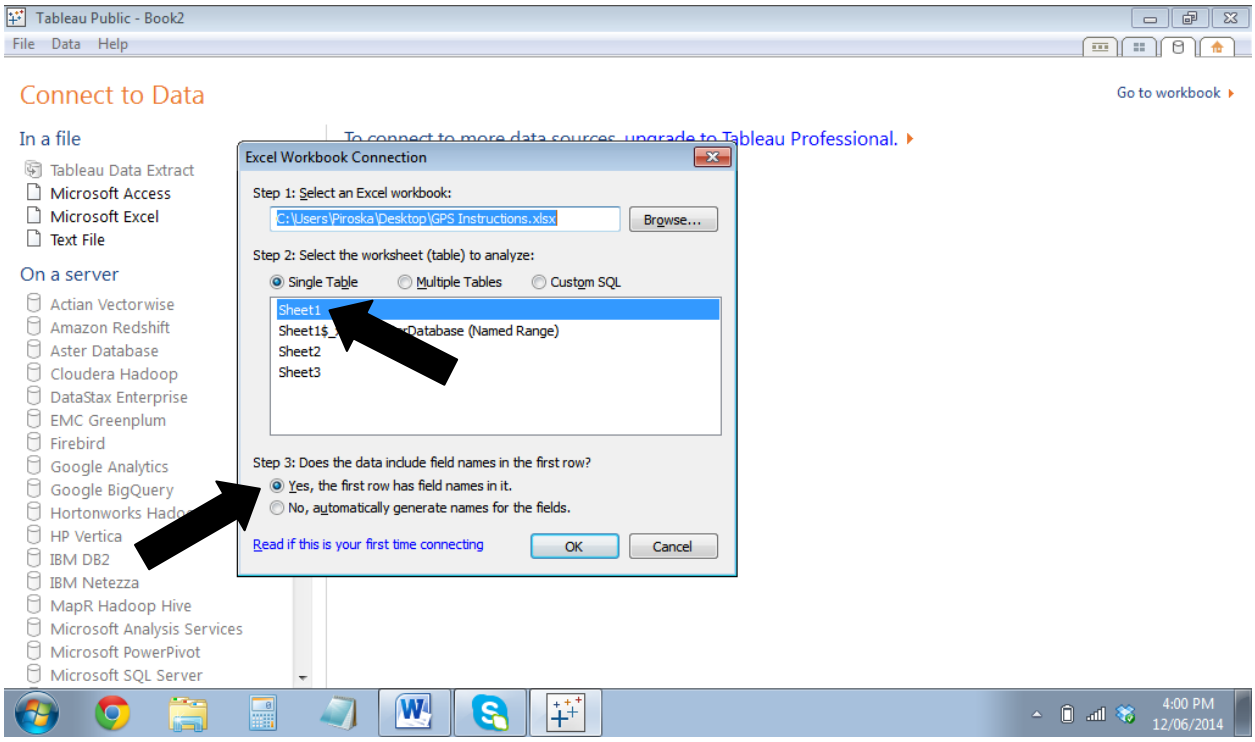
Step 3: Choose Microsoft Excel



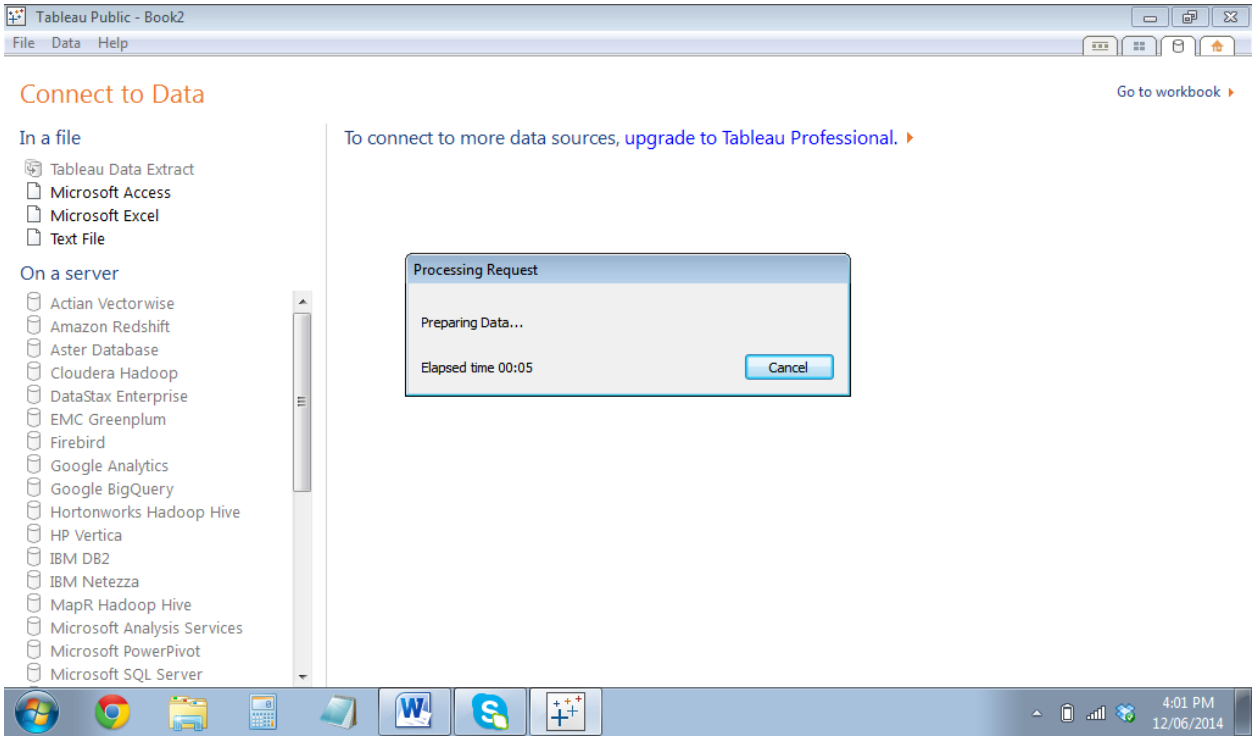
Step 4: Open the Excel file that you prepared earlier



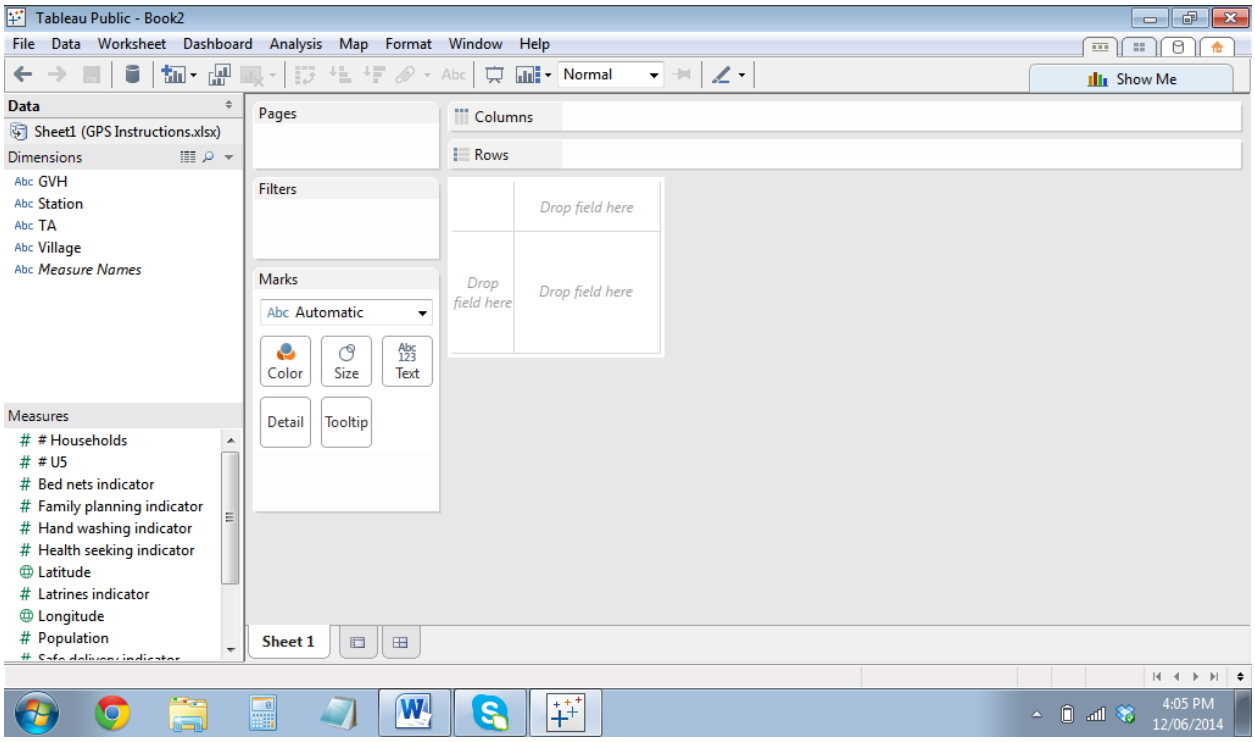
Step 5: Choose Sheet 1 and tick the box that says “Yes, the first row has field names in it.” Click OK.



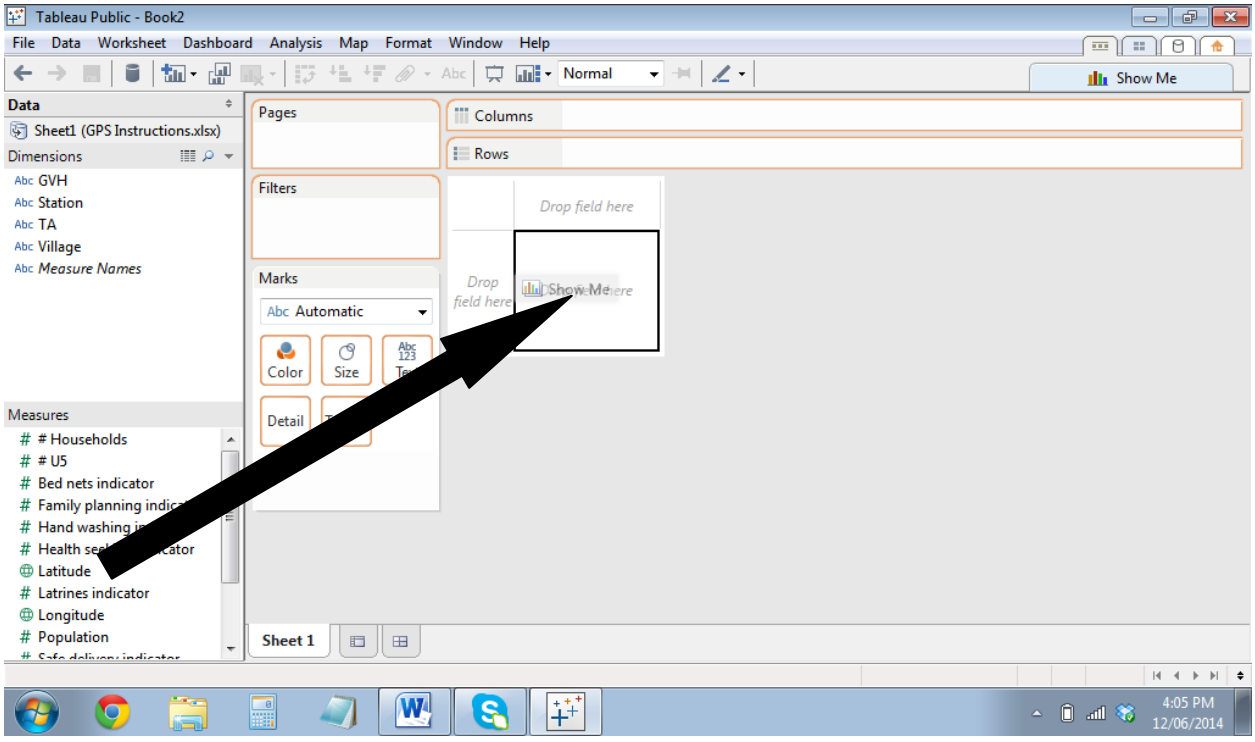
Step 6: You will see a message saying that it is preparing the data. This may take several minutes. Wait until it is finished.



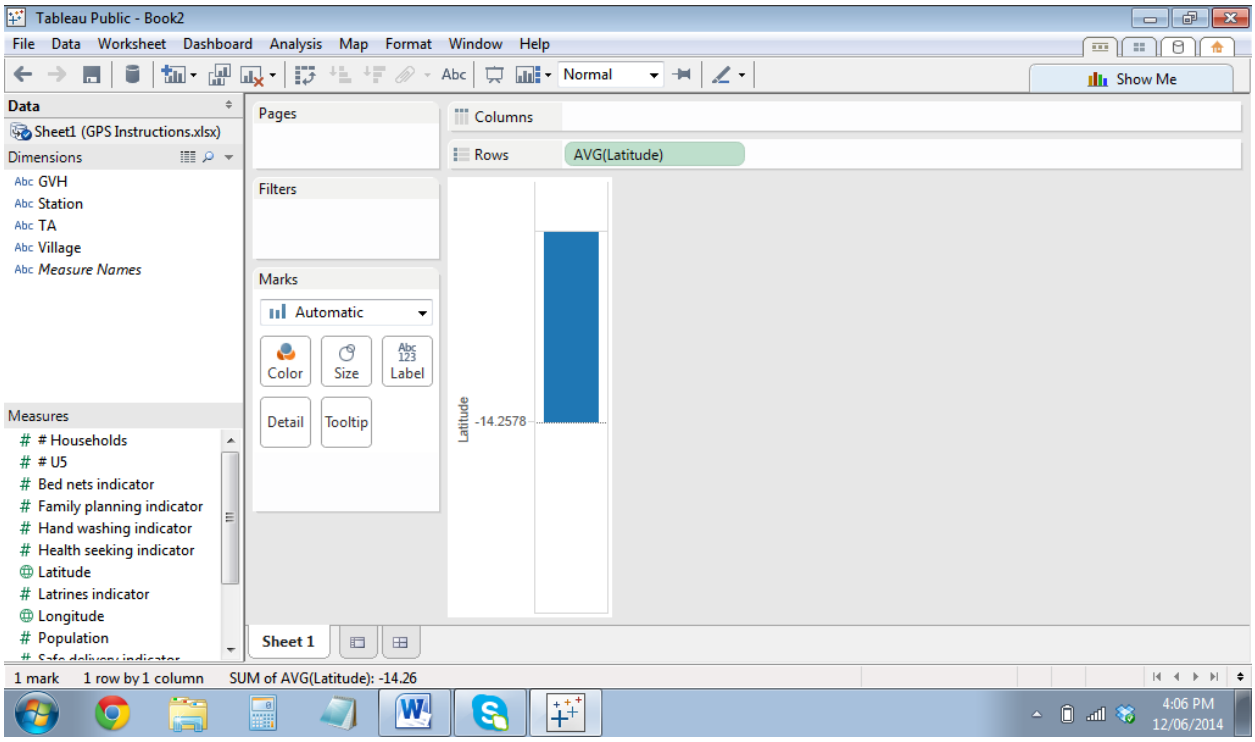
Step 7: When it has finished processing you should see a screen like this. Your column names will appear in a list down the left hand side.



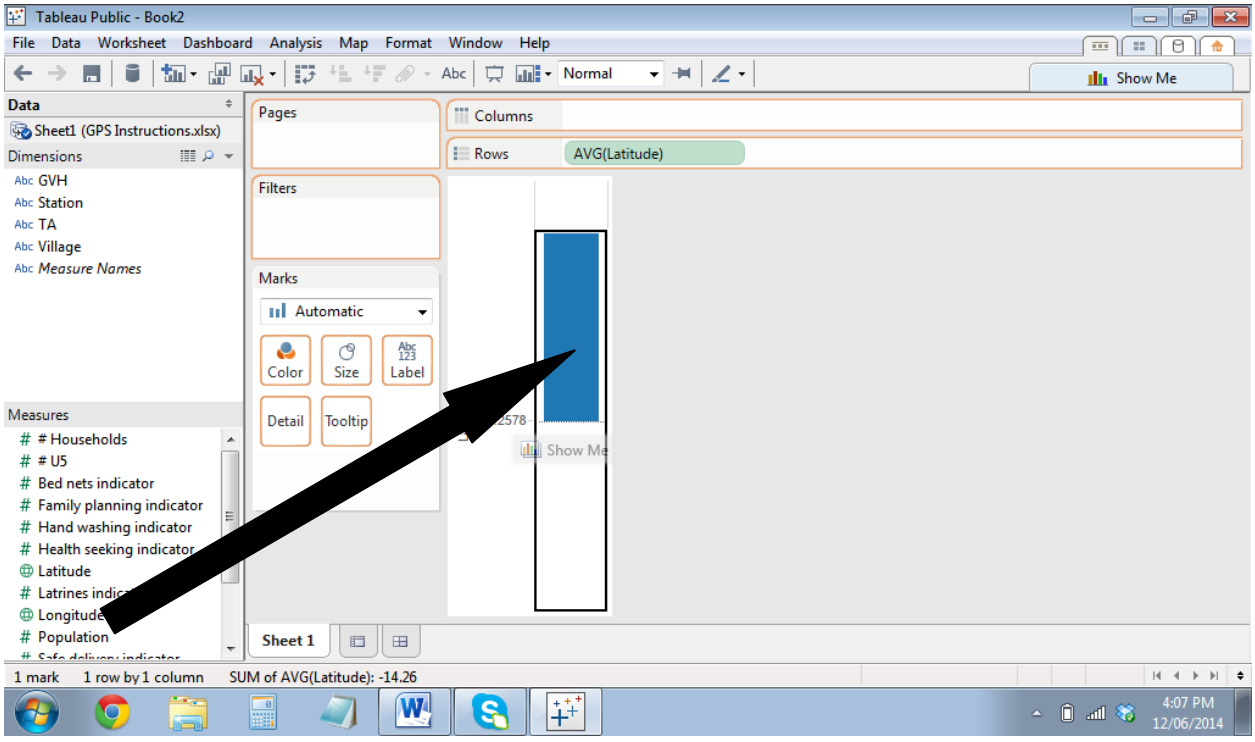
Step 8: Drag the Latitude into the middle box that says “Drop field here”



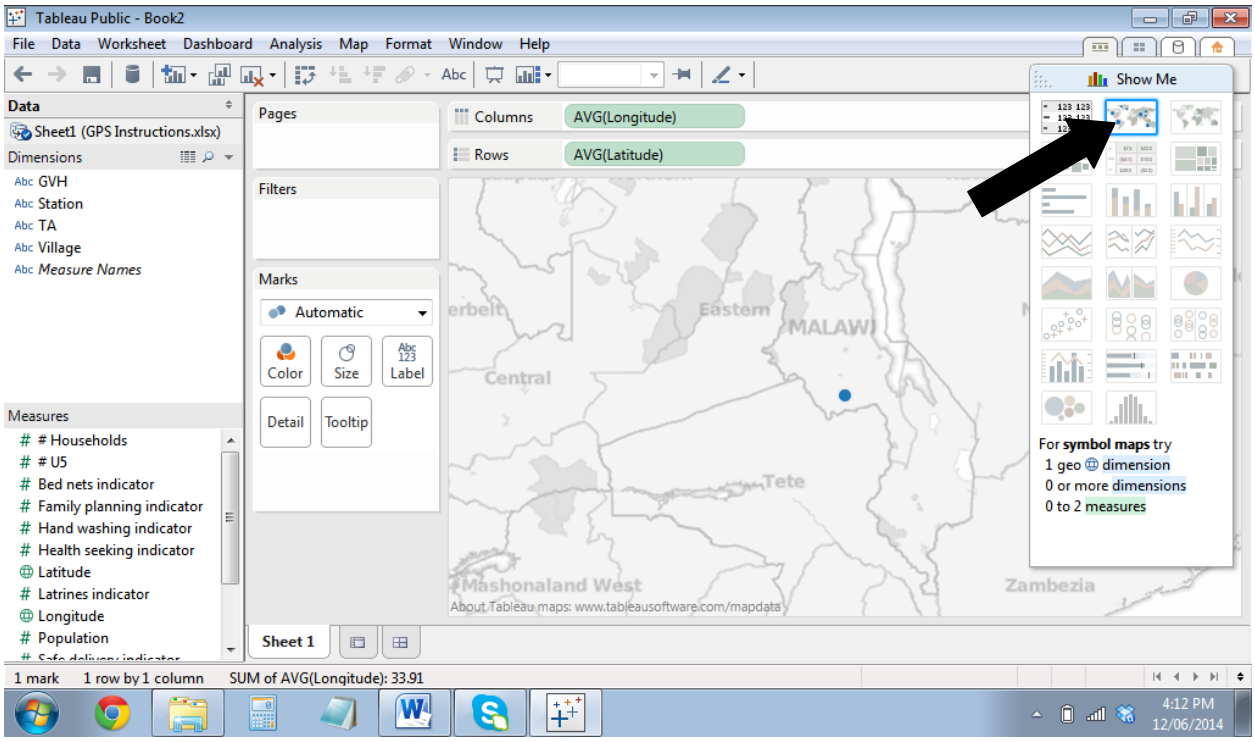
Step 9: The screen will now look like this.



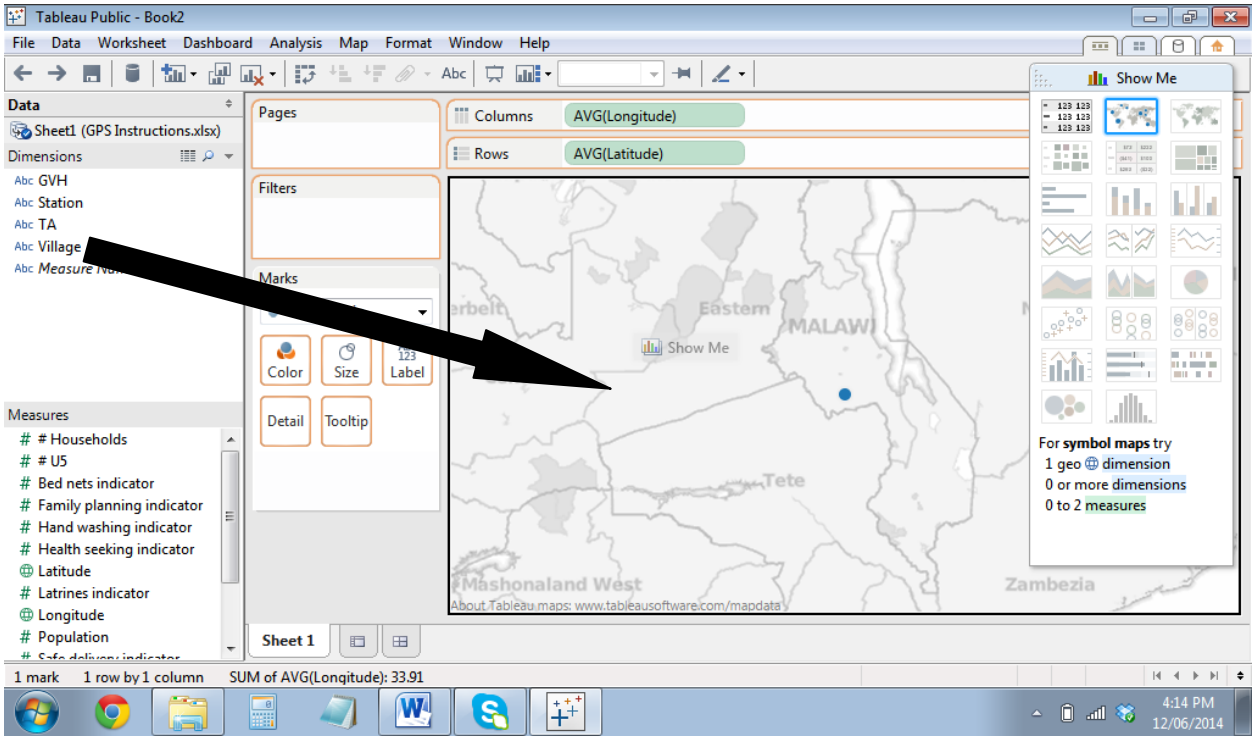
Step 10: Drag the Longitude into middle of the chart. Make sure it is right in the middle, and not on the top or side.



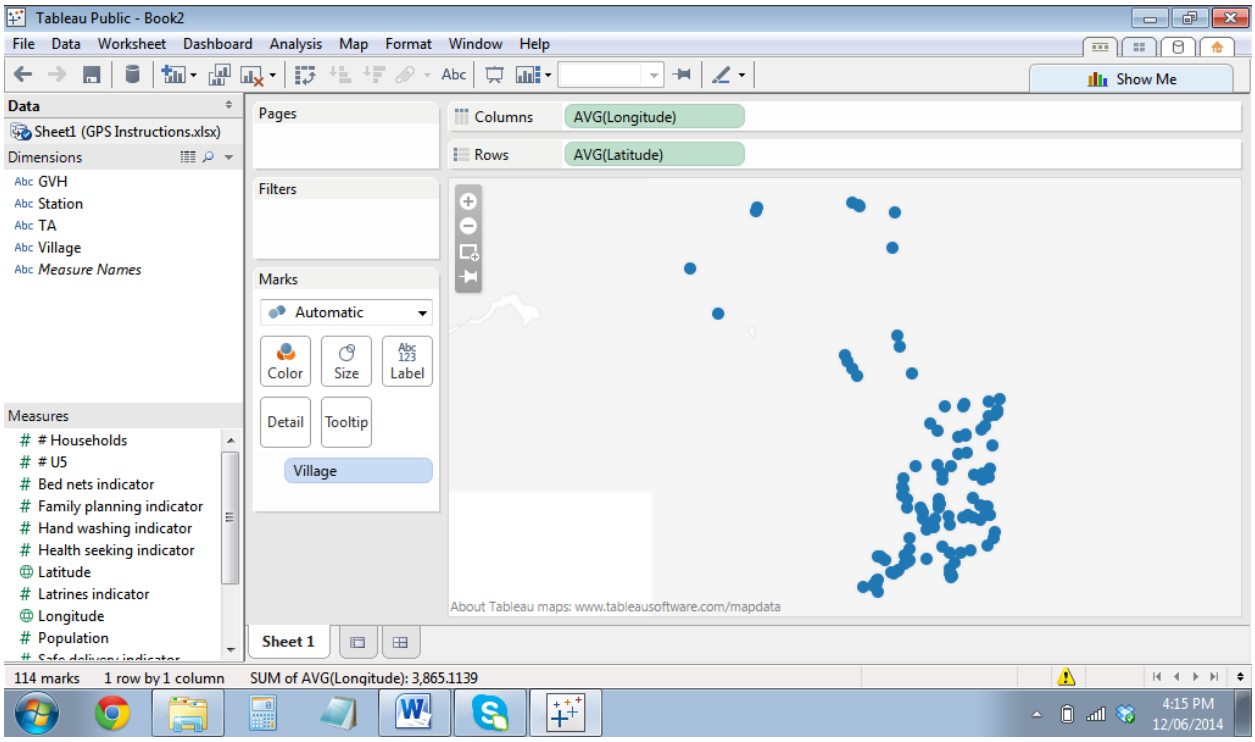
Step 11: The screen will now look like this. If it does not look like this then click on the Show Me bar and choose the map icon with dots.



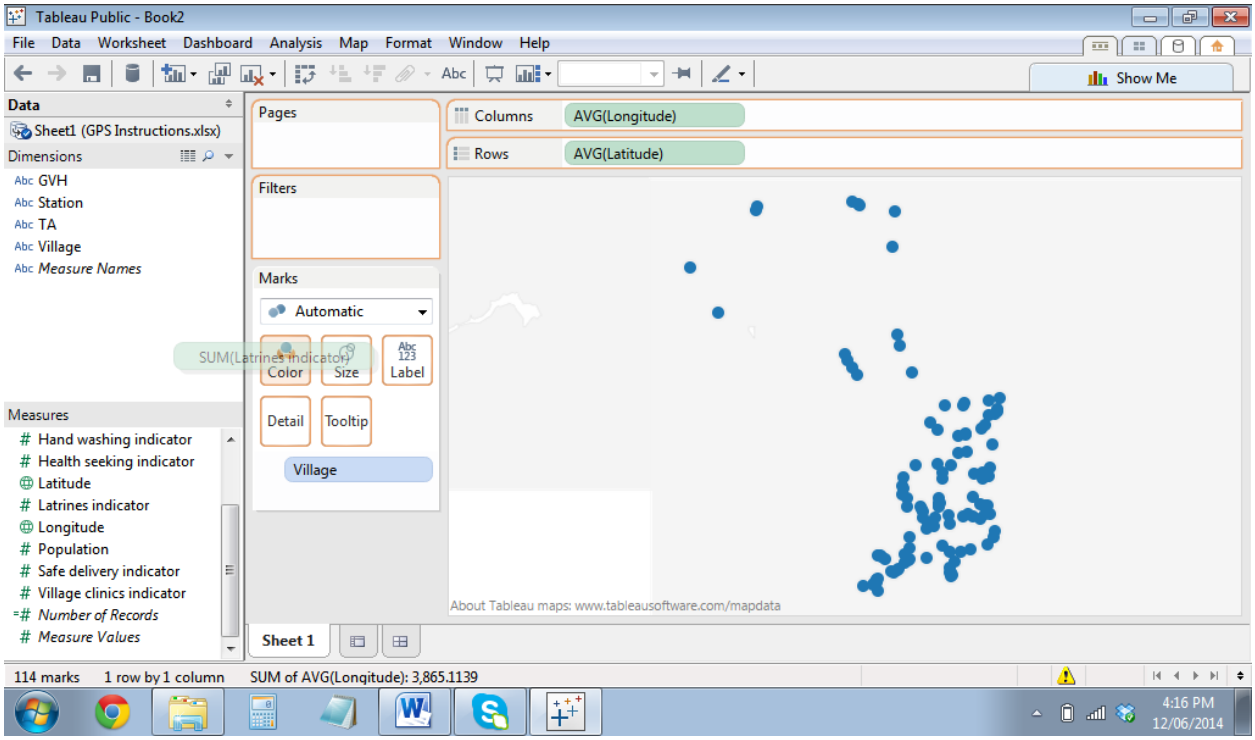
Step 12: Drag Village name onto the map.



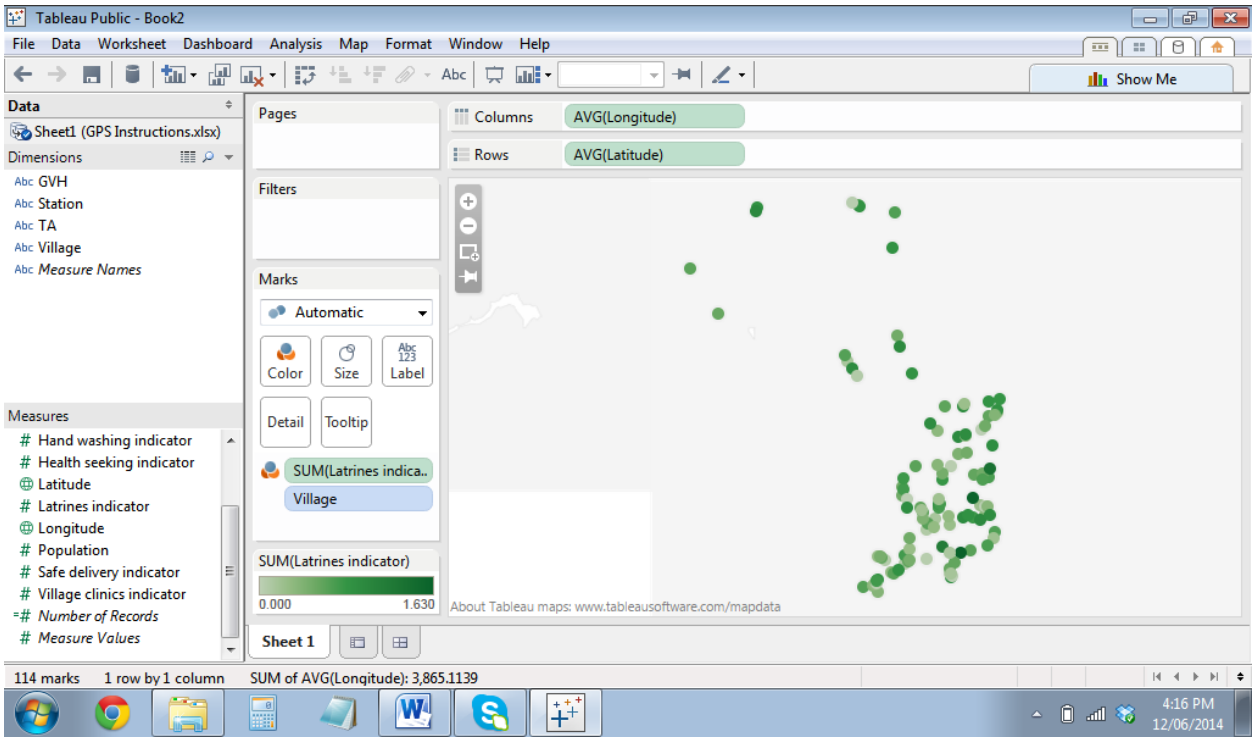
Step 13: The screen should now look like this. To zoom in and out use the + and - buttons in the top left corner of the map. If you don't see the map behind the dots then it is probably because your internet is too slow. If you see dots in the wrong place (like outside Malawi, or in the ocean) then go back to the spreadsheet and check that there are no errors in the latitudes and longitudes.



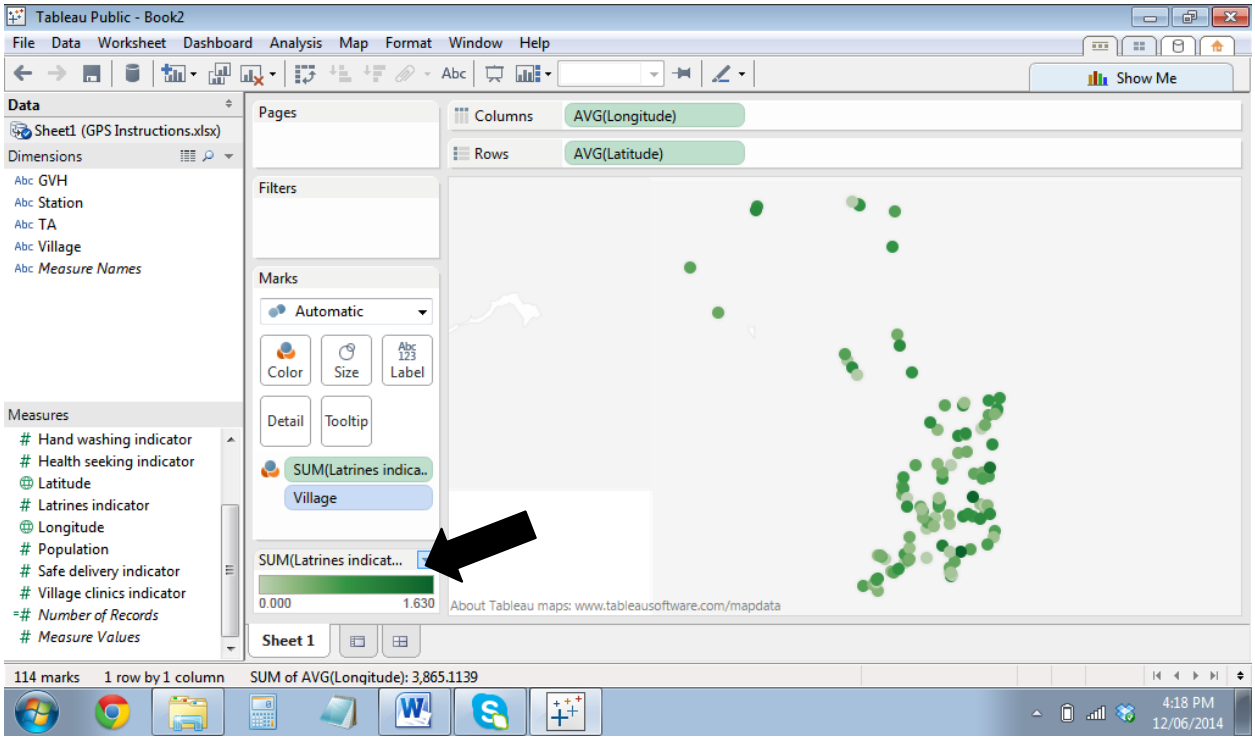
Step 14: To change the colour of the dots, drag the item you want to colour by (in this example it is the latrine indicator) onto the Colour icon.



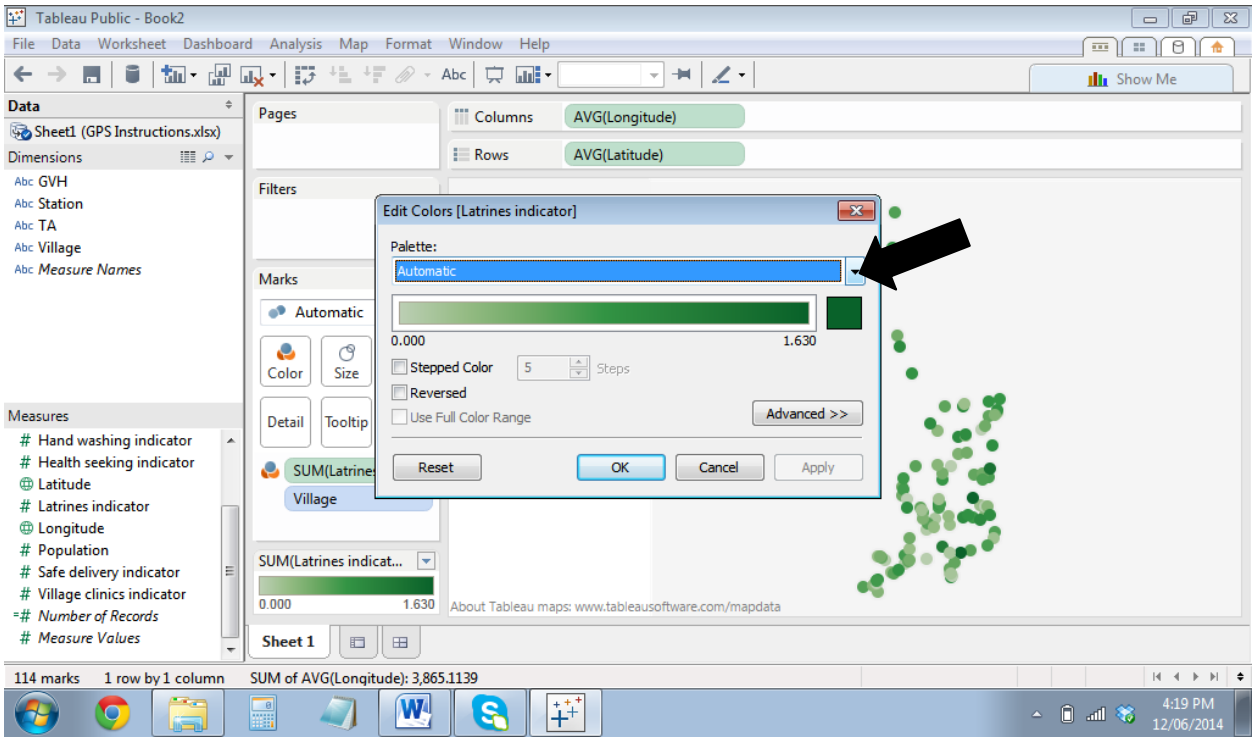
Villages with high latrine coverage are now dark green and low latrine coverage is light green.



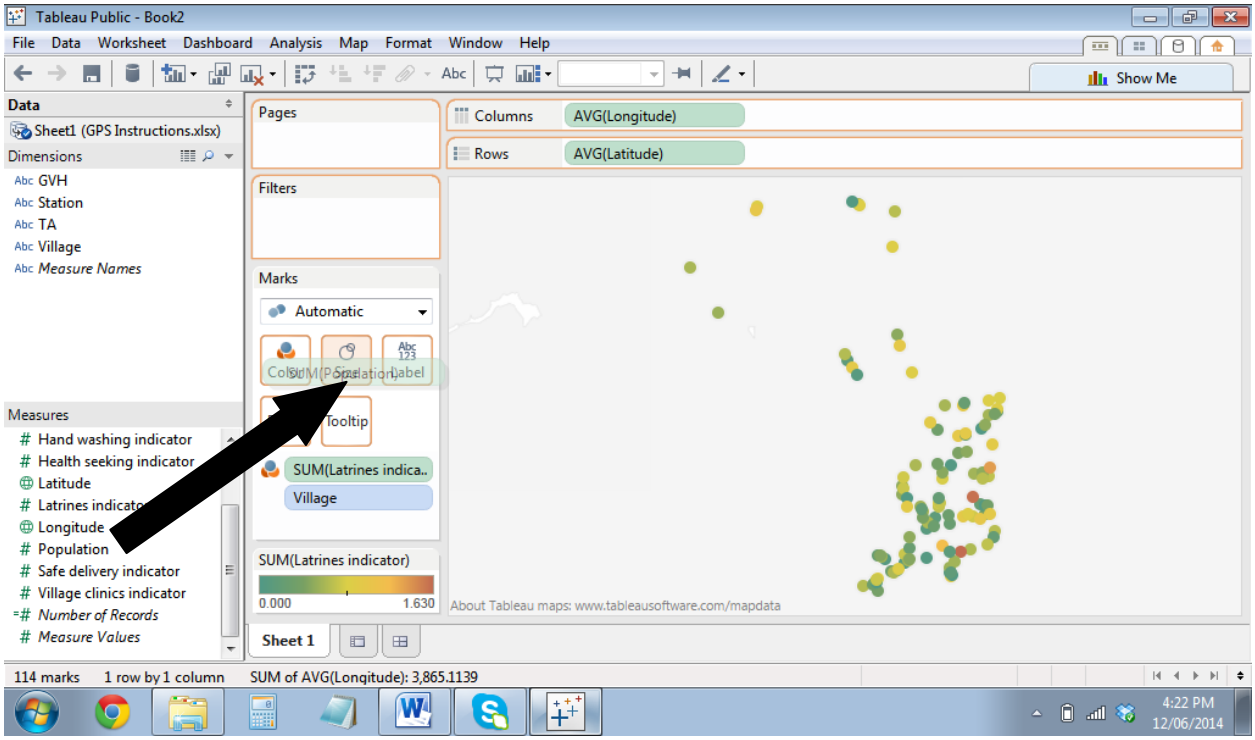
Step 15: To change the colour scheme click on the small arrow next to the colour scheme title. Choose Edit Colours.



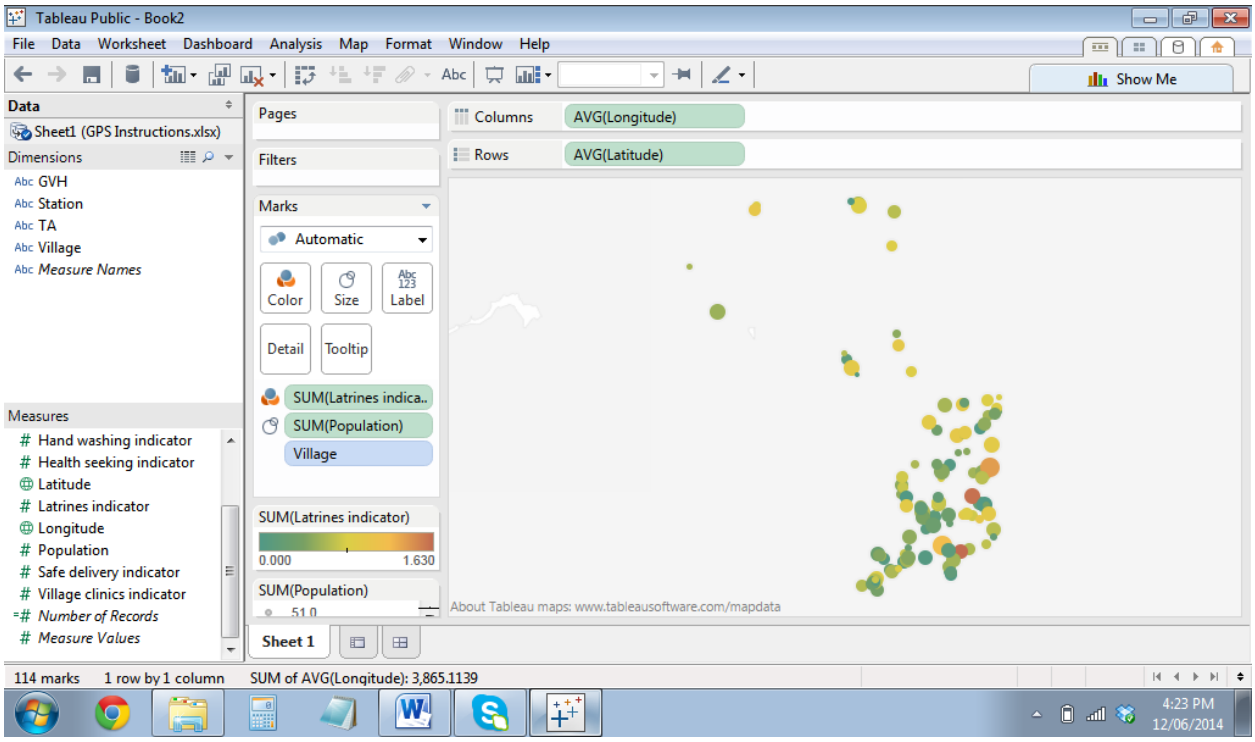
To choose a new colour scheme click on the small arrow next to Automatic.



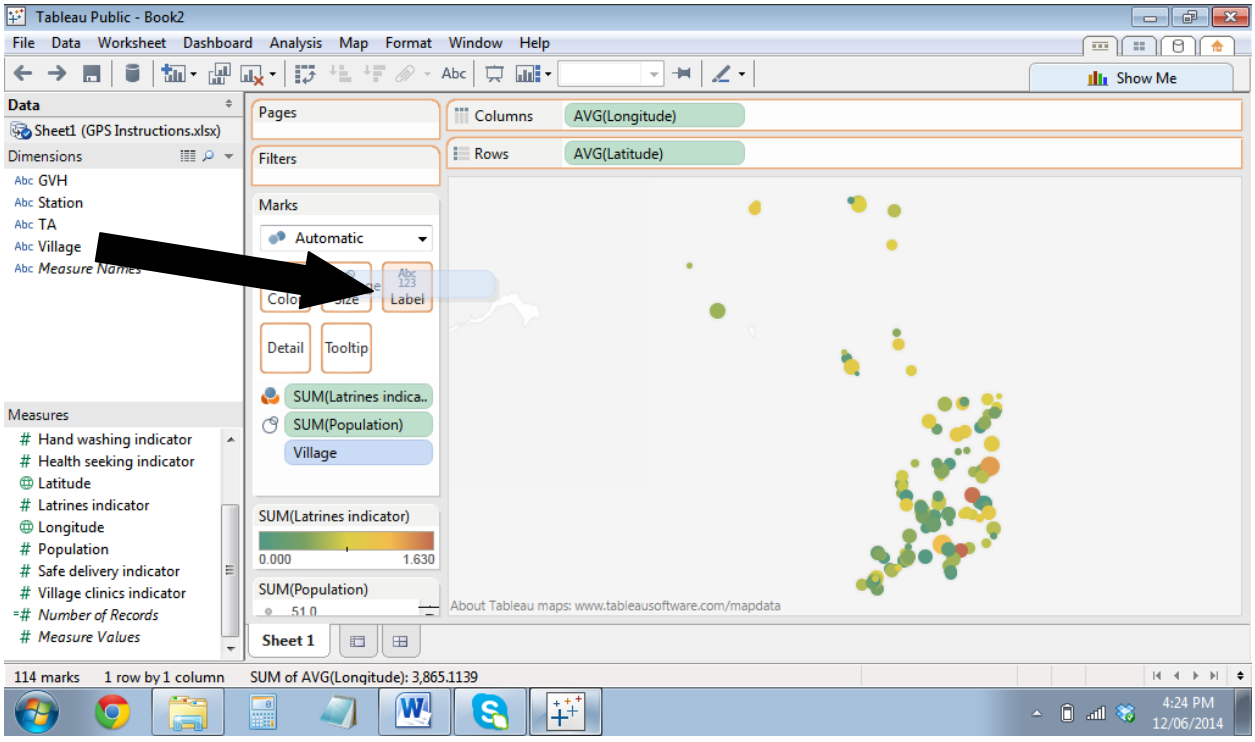
Step 16: To change the size of the dots, drag the item you want to use for the size (in this example it is the Population) onto the Size icon.

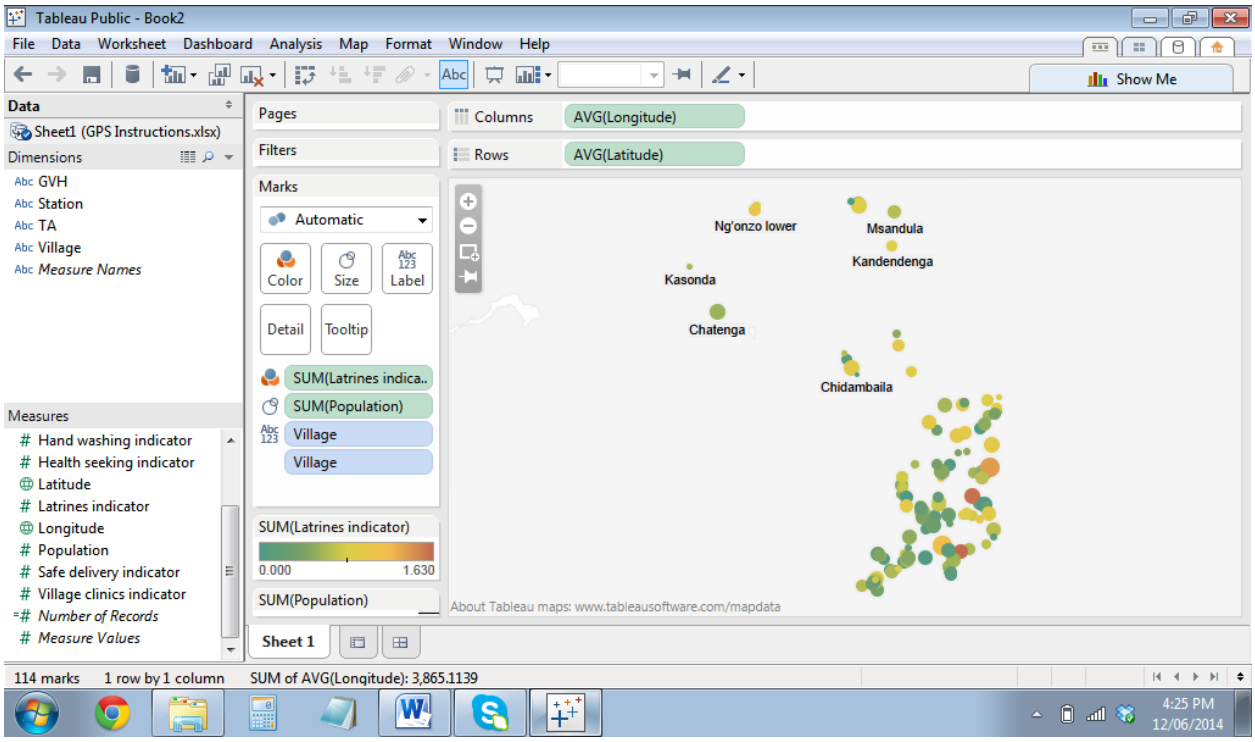


Villages with a higher population will now appear as bigger dots.

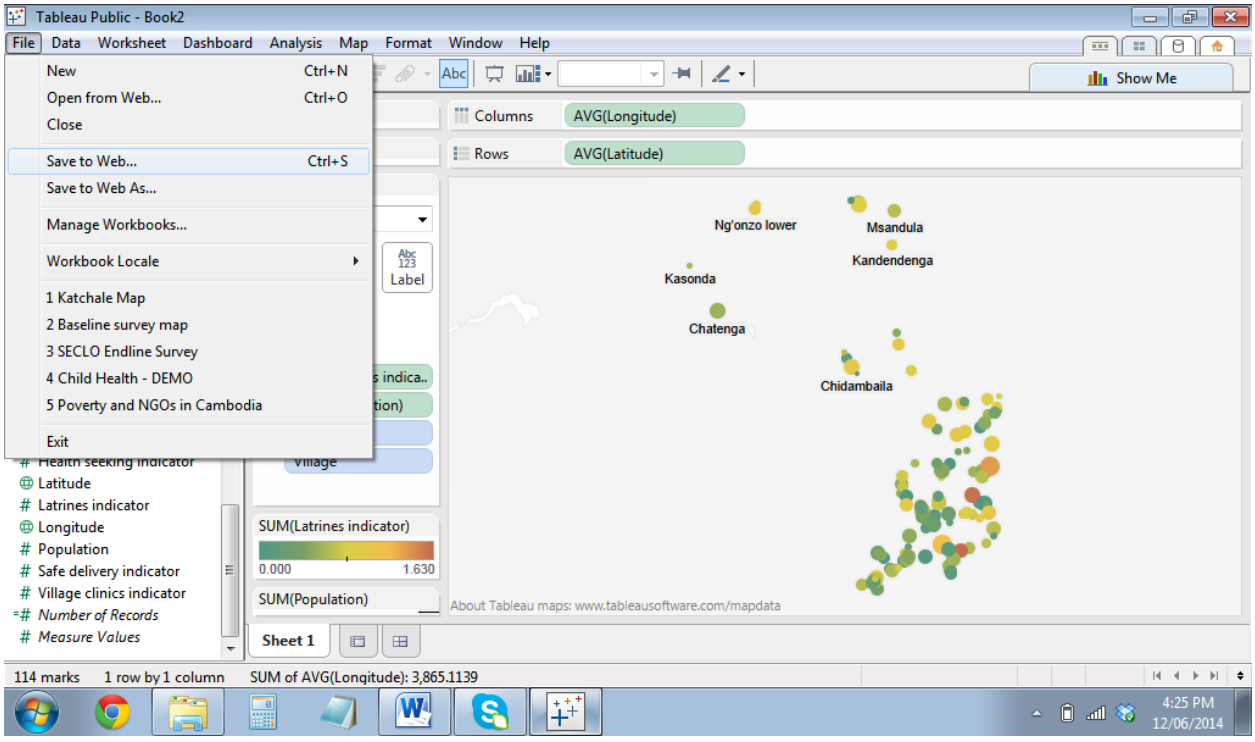


Step 17: To show the names of villages drag Village name onto the Label icon.

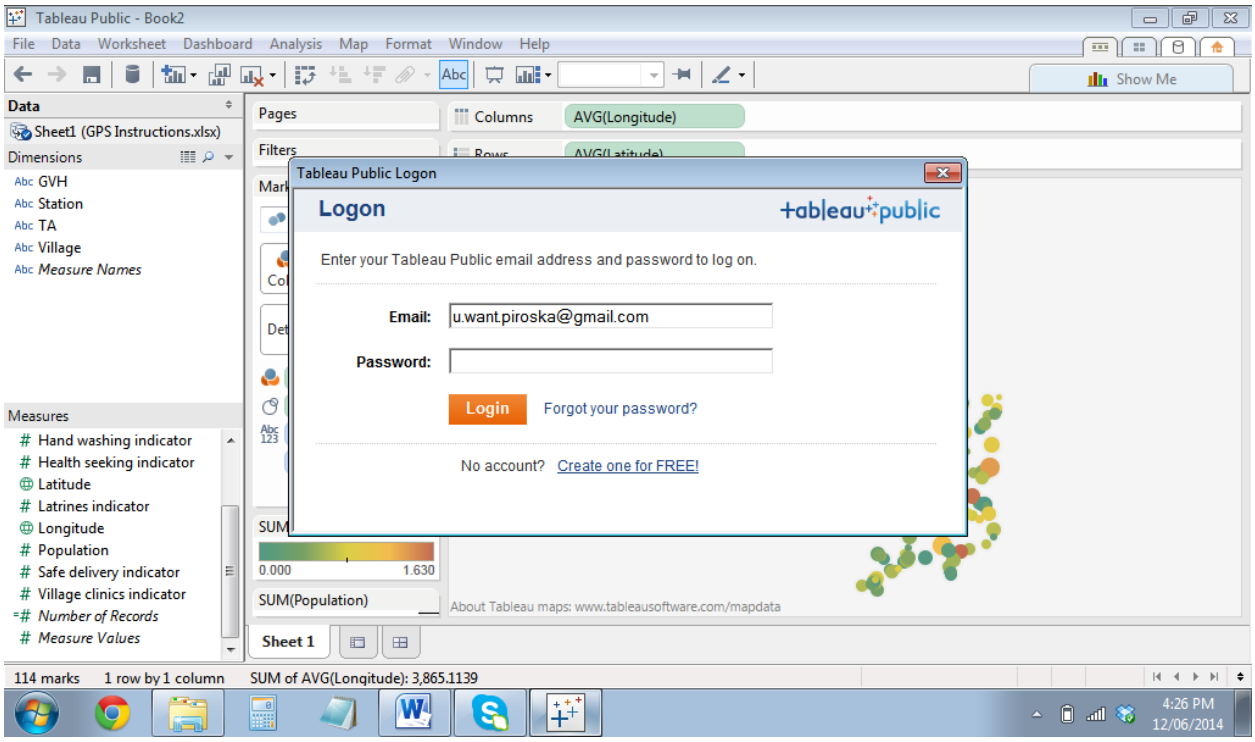




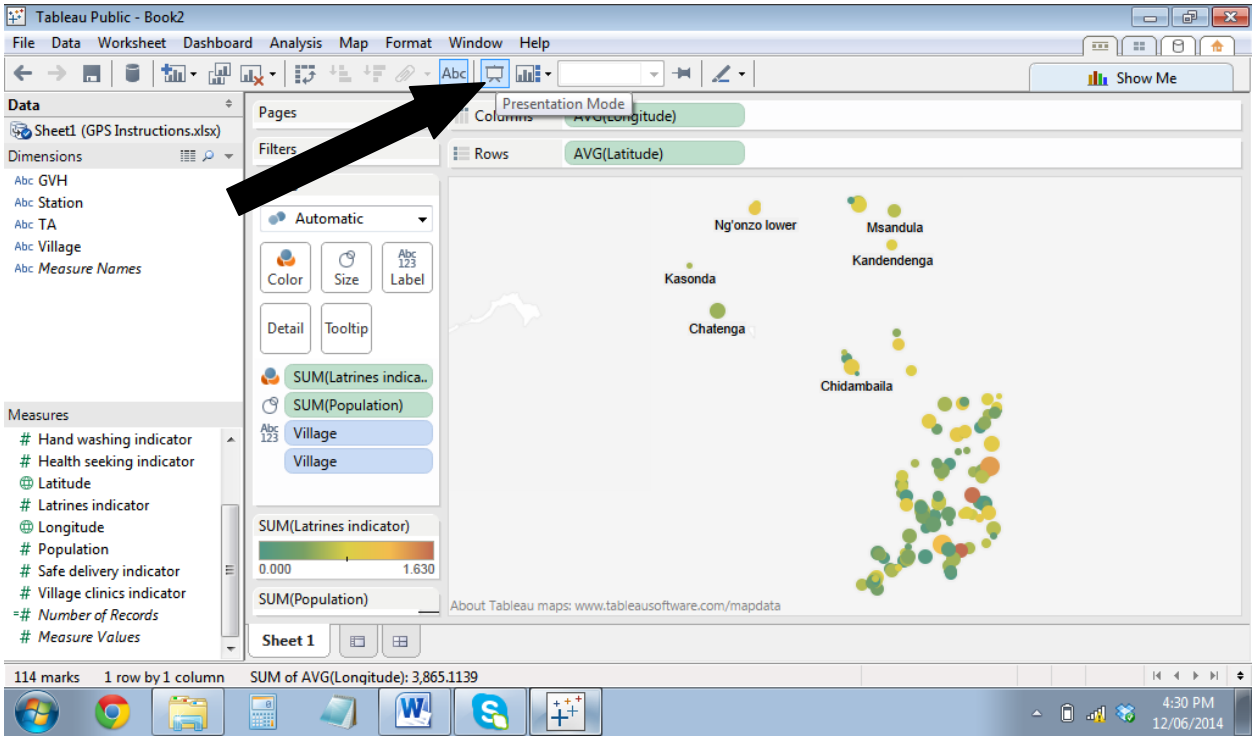
Step 18: To save the map click File > Save to Web



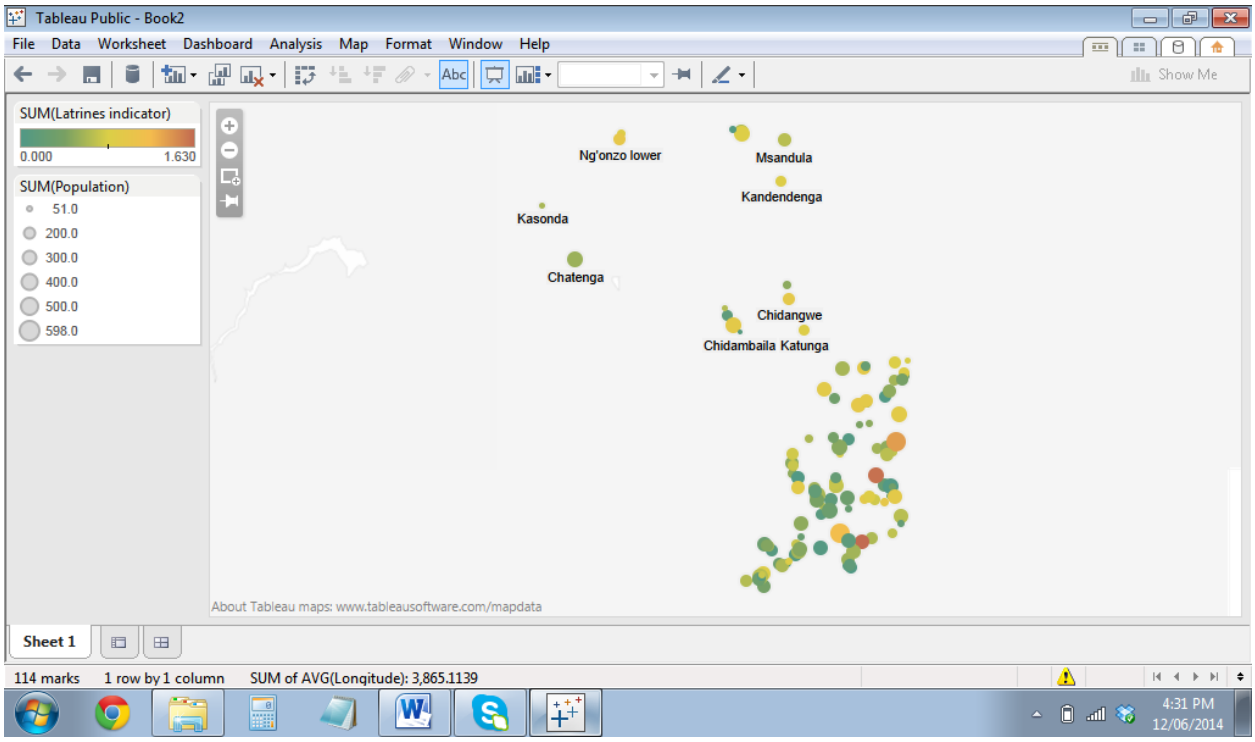
Step 19: To save it you will need to log on to Tableu Public with your email address and password. If you do not have an account then click on Create one for FREE!. Once you have setup the account then log in to save the map.



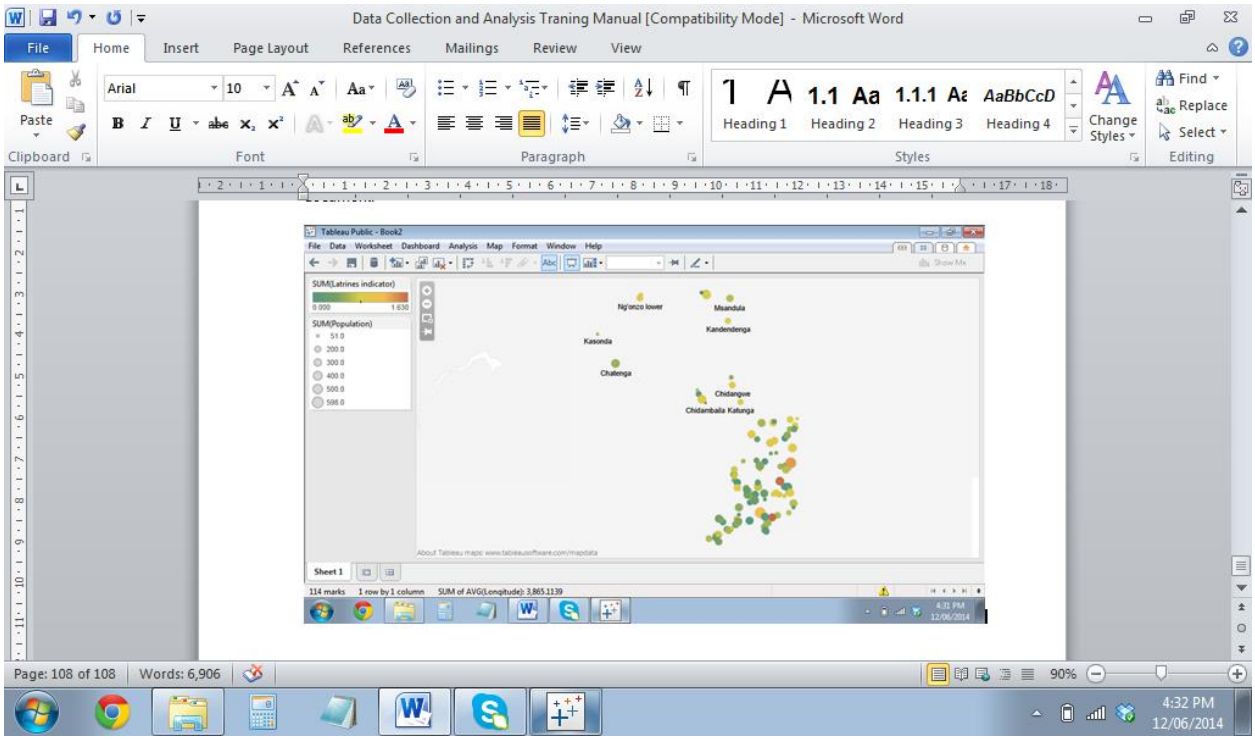
Step 20: To insert the map into a Word document, click on the presentation icon.



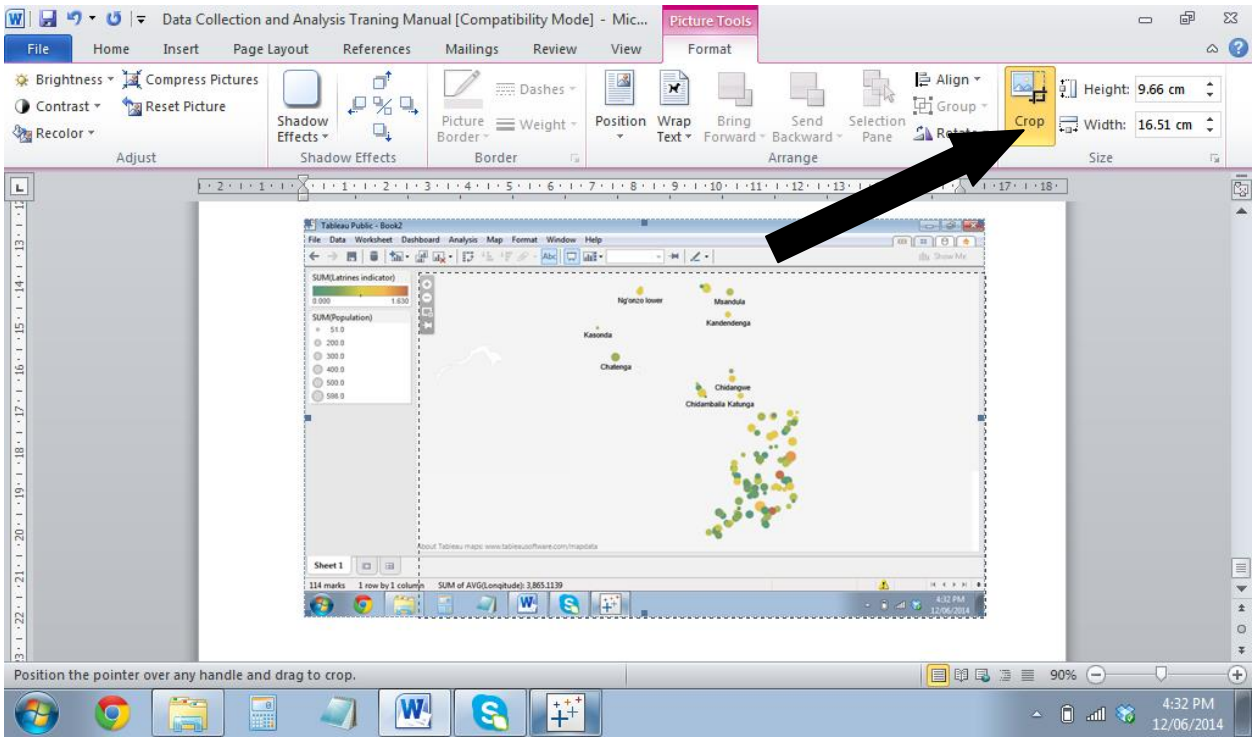
Step 21: The map will appear in large size. Click the PrtScr button on your keyboard to take copy of the screen.

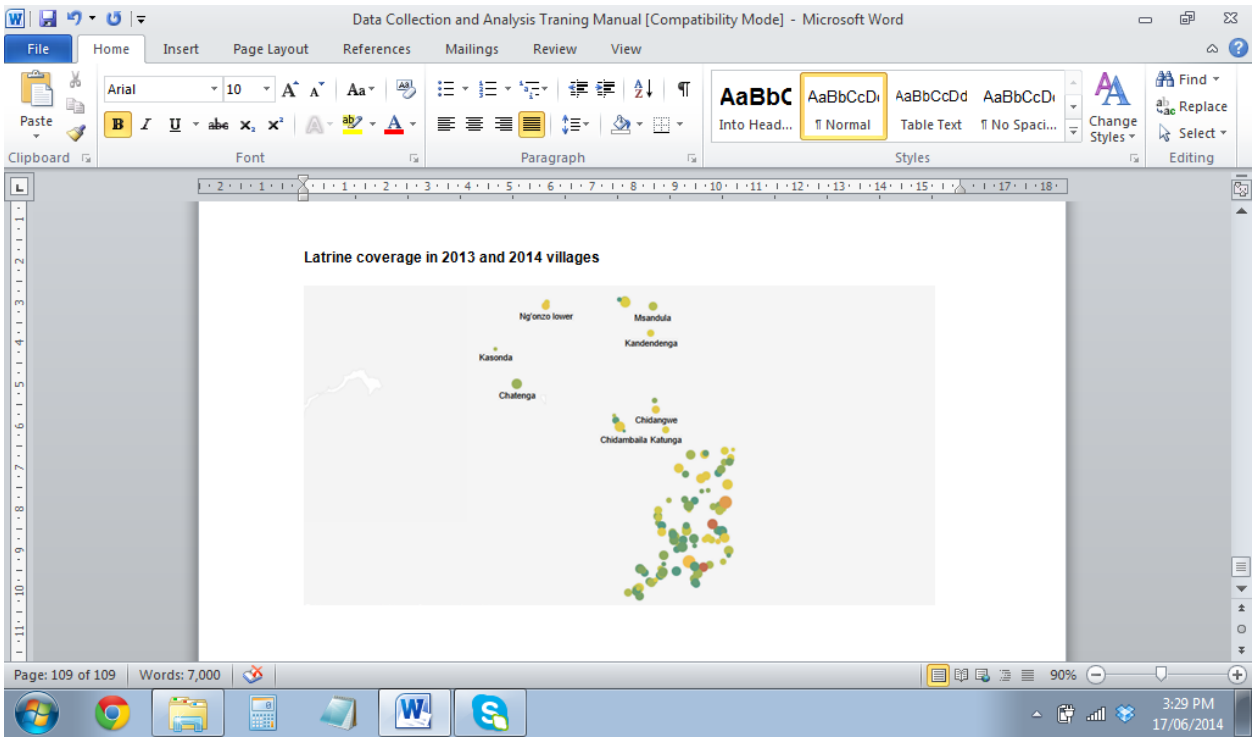


Step 22: Go to the Word document and click paste. A copy of the screen will appear in the Word document.



Step 23: Use the crop feature in Word to remove the unnecessary parts of the image.


Data Collection and Analysis Traning Manual [Compatibility Mode] - Mic...
Picture Tools
File
Home
Insert
Page Layout
References
Mailings
Review
View
Format
Brightness
Contrast
Recolor
Compress Pictures
Reset Picture
Adjust
Shadow Effects
Shadow Effects
Picture Border
Dashes
Weight
Border
Position
Wrap Text
Bring Forward
Send Backward
Selection Pane
Align
Group
Crop
Height: 9.66 cm
Width: 16.51 cm
Arrange
Size
Ng'onzo lower
Msandula
Kandendenga
Kasonda
Chatenga
Chidangwe
Chidambalila Katunga
Position the pointer over any handle and drag to crop.
90%
4:32 PM
12/06/2014


Data Collection and Analysis Traning Manual [Compatibility Mode] - Microsoft Word
File
Home
Insert
Page Layout
References
Mailings
Review
View
Paste
Clipboard
Arial
10
Font
Paragraph
Into Head...
Normal
Table Text
No Spaci...
Styles
Change Styles
Find
Replace
Select
Editing
Latrine coverage in 2013 and 2014 villages
Ng'onzo lower
Msandula
Kandendenga
Kasonda
Chatenga
Chidangwe
Chidambalila Katunga
Page: 109 of 109
Words: 7,000
90%
3:29 PM
17/06/2014