



Tutorial on risk factor identification in epidemiological studies using epi-info



Contents

1. Defining variables to be used in the study- outcome, exposure, predictors, potential confounders	2
2. Designing questionnaires for risk factor identification	Error! Bookmark not defined.
3. Development of Questionnaire for Identification of Risk Factors Using Epi-Info	3
4. Entering of Field Collected Questionnaire Data in Epi-Info	13
5. Analysis of Epidemiological Data Using Visual Dashboard	18
5. Exercises	25

1. Defining variables to be used in the study- outcome, exposure, predictors, potential confounders

Variable: A variable is a characteristic that varies from one observation to the other and can be measured. For example sex is a variable because it can be either male or female

Outcome variable and exposure variable: In epidemiological studies, occurrence of a disease is dependent on environmental and agent factors. Here the occurrence of a disease is *outcome variable* and the factors responsible for occurrences of the disease are called as *exposure variables*

Potential confounding variables: Confounding can occur when an association between a risk factor and disease can be explained by a factor associated with both disease and risk factor. To aid identification of a confounder a causal diagram is helpful.

2. Development of Questionnaire for Identification of Risk Factors Using Epi-Info

The objective of this practical exercise is to develop a questionnaire using *Epi-Info*. The questionnaire can be developed directly in *Epi-Info* or it can be created in word and once the data is collected from field, it can be created in *Epi-Info*.

Requirement for this practical

1. You should clearly able to define your research questions and purpose of using questionnaire
2. You should have a proper study design before you start working on your questionnaire
3. Your sampling design and methodology should be well defined
4. Your sample size should be estimated before collecting the data using questionnaire
5. It is always desirable to arrive at probable risk factors for a disease for which the questionnaire is planned to be developed
6. You should be clear whether the questionnaire will be based on interview or will be sent by mail
7. It is advisable to ask closed ended questions to the farmers/respondent for risk factor identification.

Learning outcome of this practical

1. You should be able to design your questionnaire using *Epi-Info*
2. You should be able to create questionnaire database in *Epi-Info* from your hard copies or word documents

Post development of questionnaire

1. Pilot testing of the questionnaire is important so that you can make any required changes before going for the proper survey in field.
2. Questionnaire survey to be carried out based on sample size and locations.

1. Creating forms using CREATE FORMS

Once installed, an **epi-info** icon is created on the desktop. Click on the **epi-info** icon and the options are reflected as shown in the Figure 1.



Figure 1

Click on the **CREATE FORMS** icon of the options. The other options; **ENTER DATA**, **CREATE MAPS**, **STATCALC**, **CLASSIC** and **VISUAL DASHBOARD** will be used in subsequent exercises.

Once you click on the **CREATE FORMS** you will see a screen with different options as shown in Figure 2. If you have created already a project in Epi-Info or if you want to use the existing examples in the software you can choose option of Open Project. Click on the **New project** as we are going to create new FORM using form designer (Figure 2).

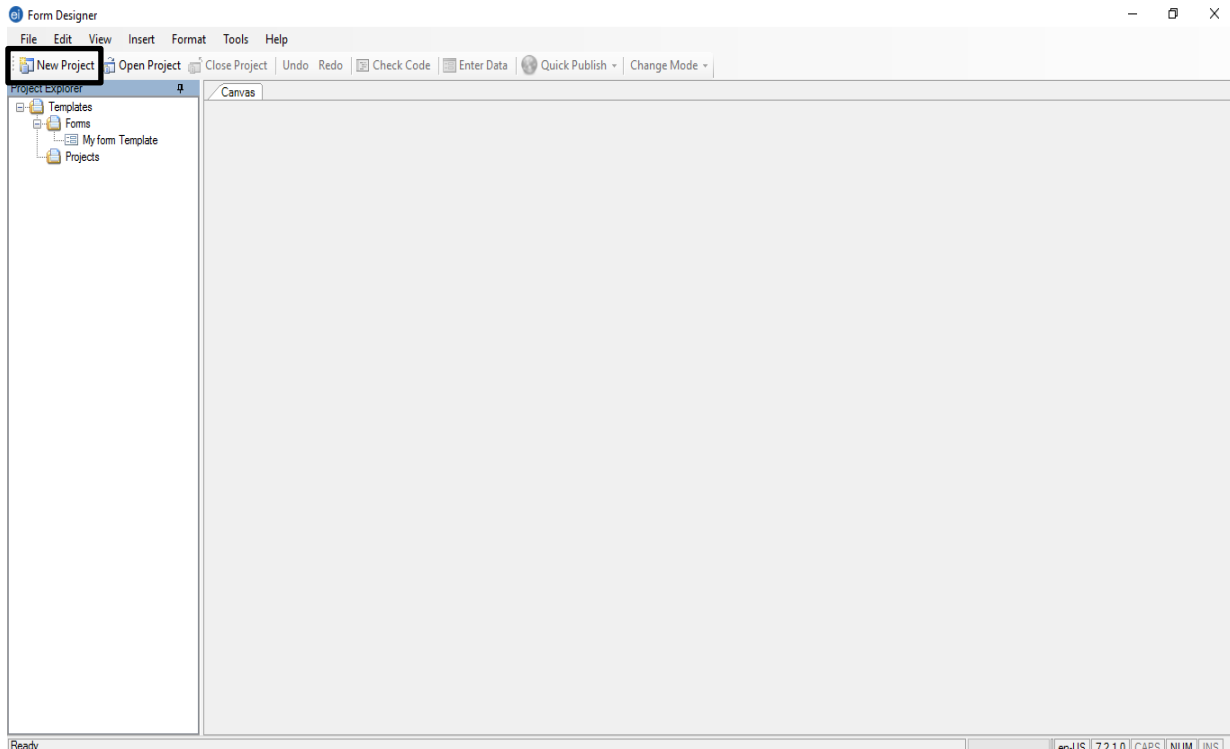


Figure 2

Once you click on the open project icon you will be asked to give name to your project. Here we will be giving **Anthrax Training** as the project name as shown in Figure 3.

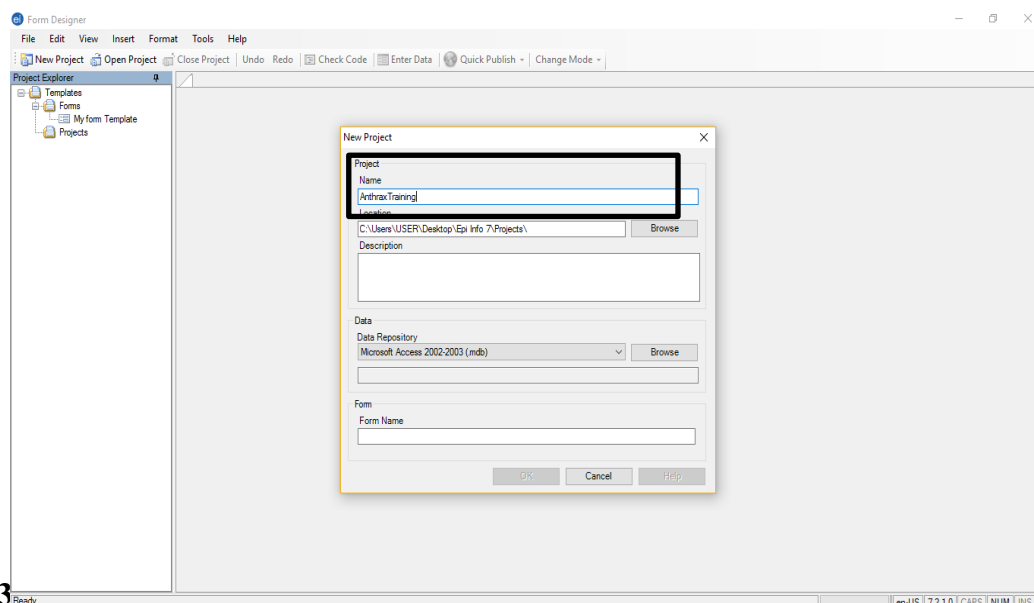


Figure 3

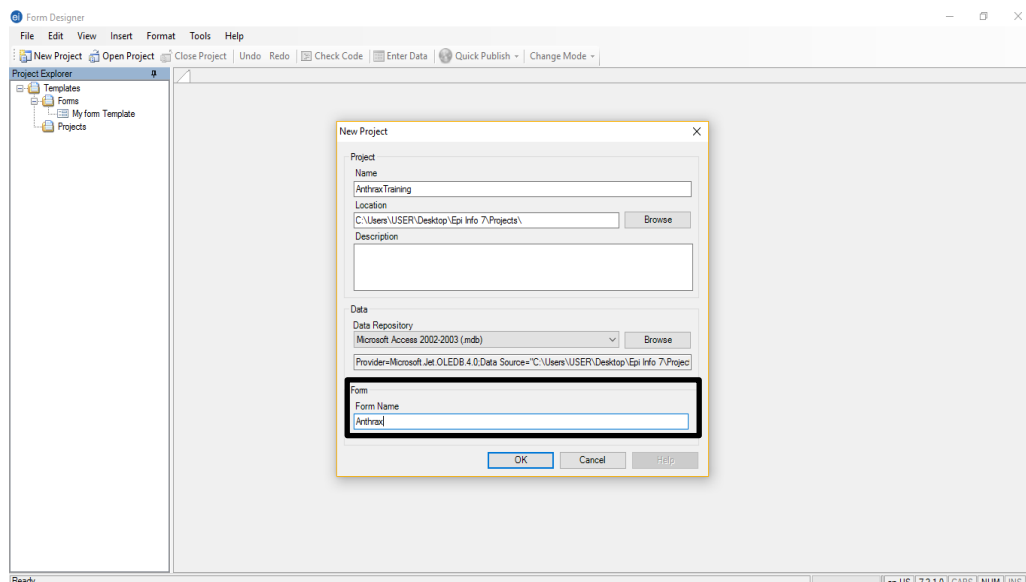


Figure 4

You need to label your form also as there is option of creating more than one form. Label your form as **Anthrax** as shown in Figure 4. A form is a database with set of questions. You can create two forms and relate them using **Relate**. A form can have many pages and each page can have a set of questions. It is important to remember the difference between page and a form when creating questionnaire database. In this exercise we are using only one form and one page, so click **OK** and a form is created.

Once you click **OK** you can observe on the left side a set of options are enabled to create database. The central blank space is also activated to add on the questions and create database (Figure 5). The options can be selected based on the type of questions you are asking and the kind of variable that will be used in the analysis.

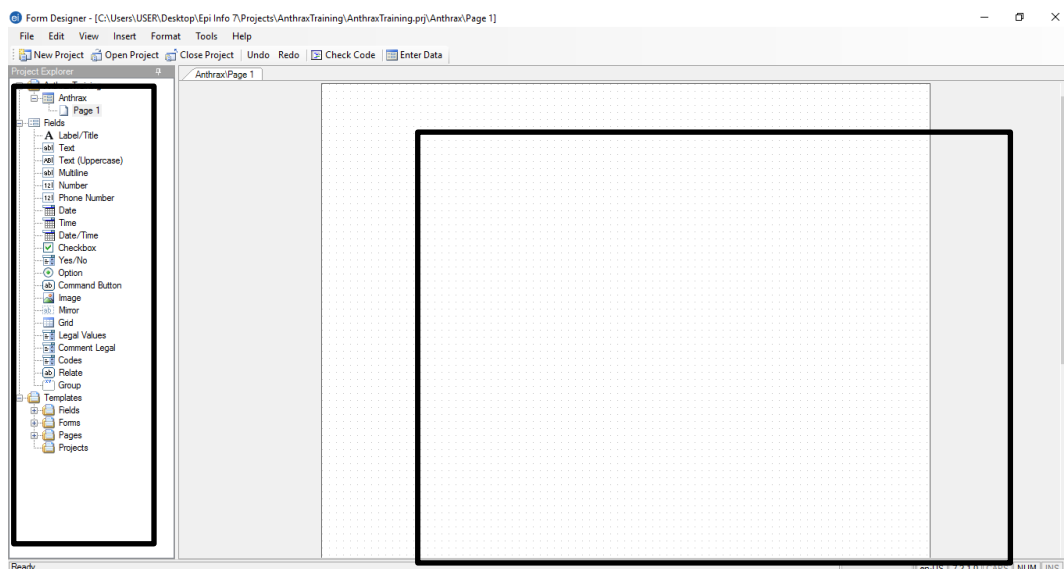


Figure 5

Label/Title: Right click on the central blank space and you will get options about the type of variable to be created (Figure 6). Click on the **Label/Title** as we will be creating a title for this questionnaire.

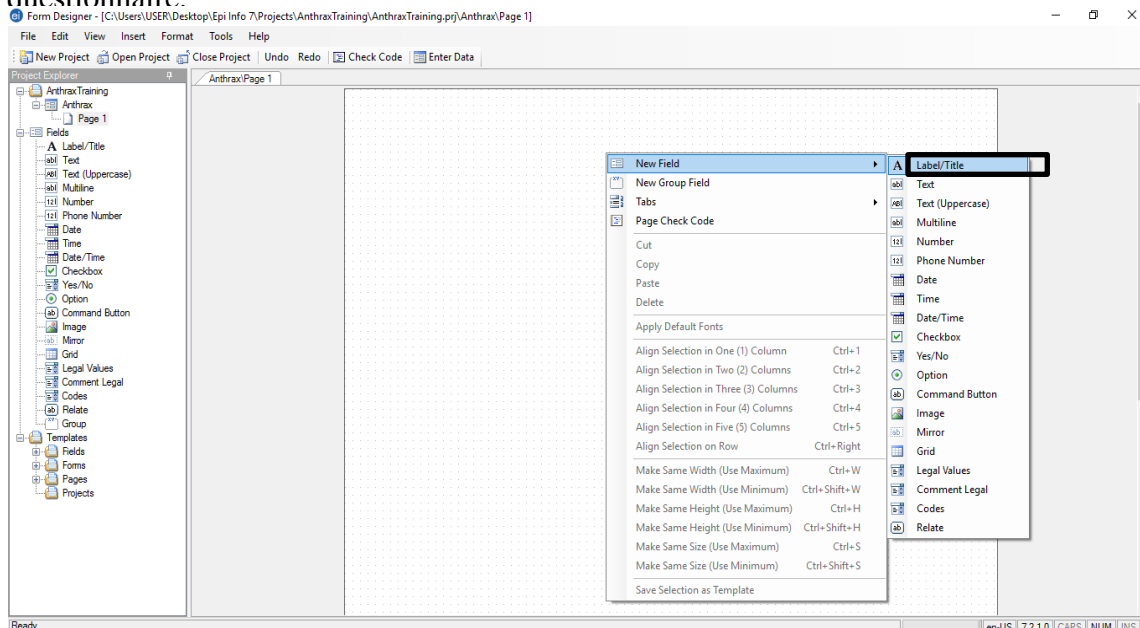


Figure 6

You will have option of typing in the blank space (Figure 7). Type the title in the blank space as **Questionnaire to identify risk factors for Anthrax in Karnataka** (Figure 8). The field name can be short as the field names will be reflected while analysis of data.

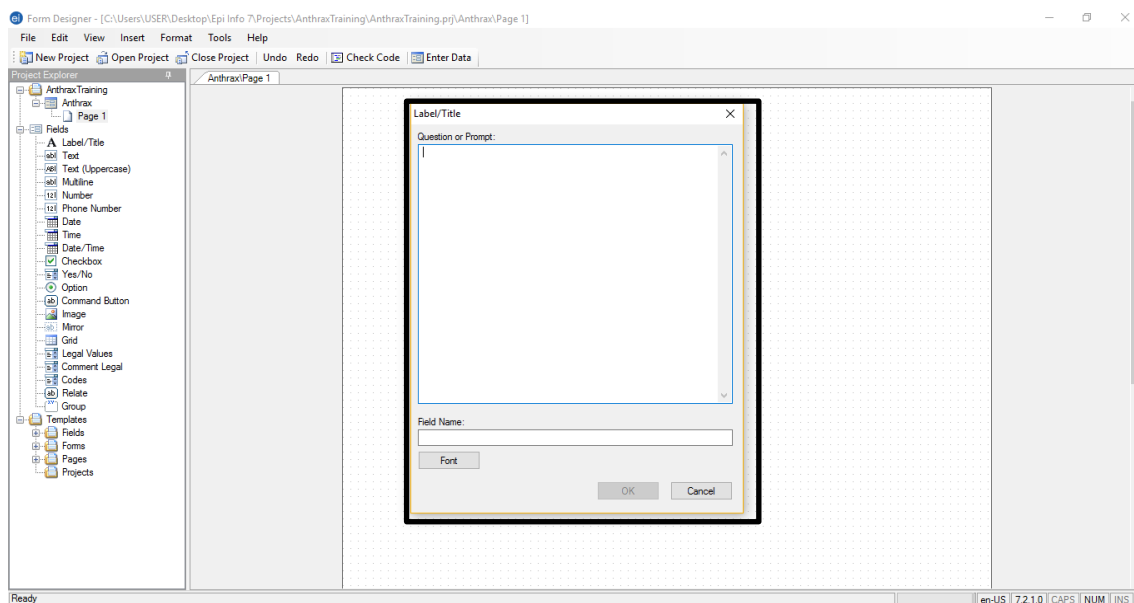


Figure 7

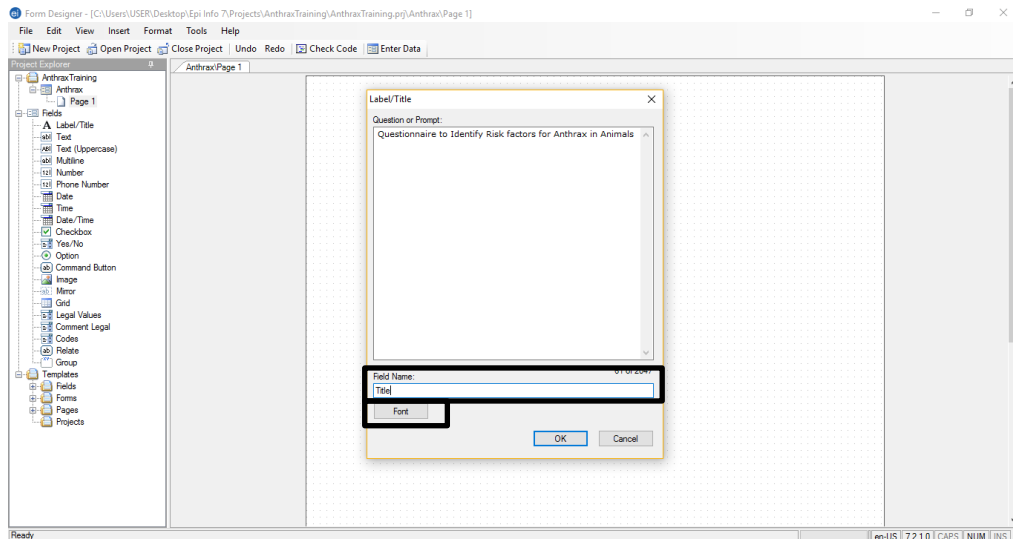


Figure 8

The font can be changed by clicking on the **Font** option.

Text, multiline and Number: Right click again to create text (**Name of the village**), multiline (**Address**) and number (**Number**) as shown in Figure 9.

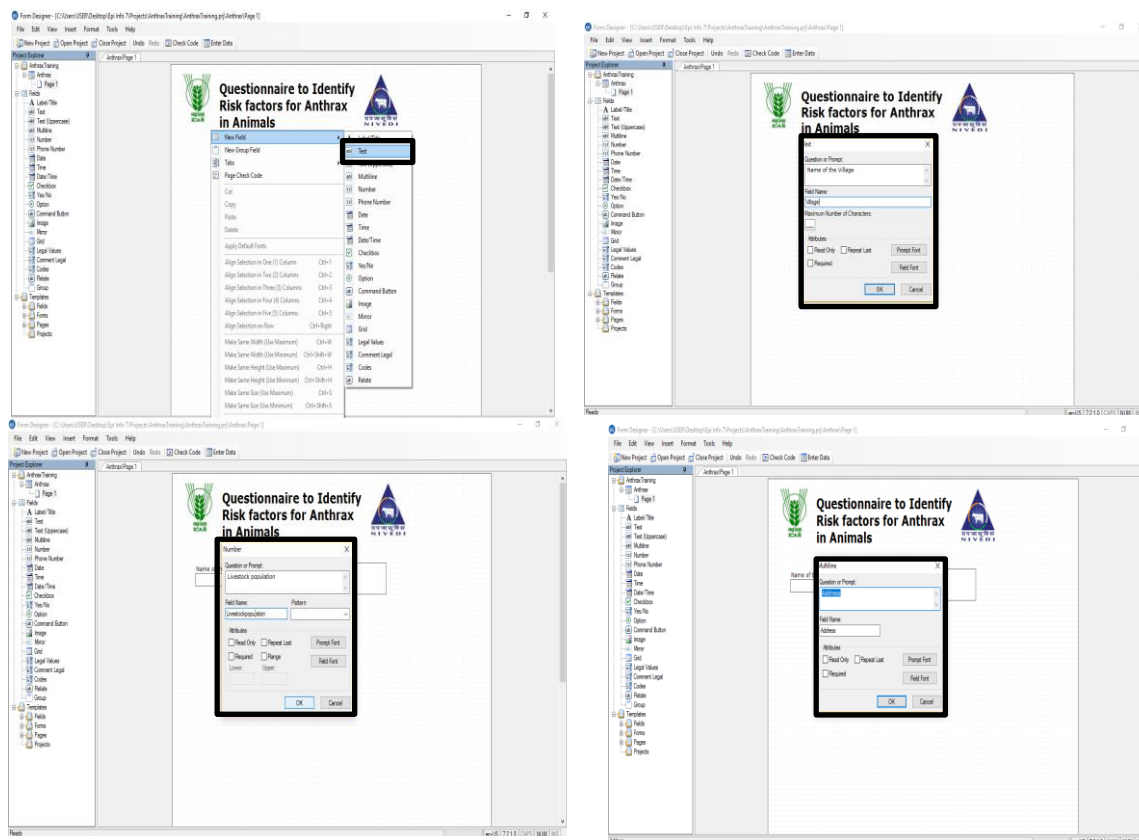


Figure 9

Checkbox: If you want to give options and it is more likely that there will be more than two options then you can create check box option as shown in Figure 10. Here you create a check box option of disposal of the carcass. So type **Discarded within farm/home premises**.

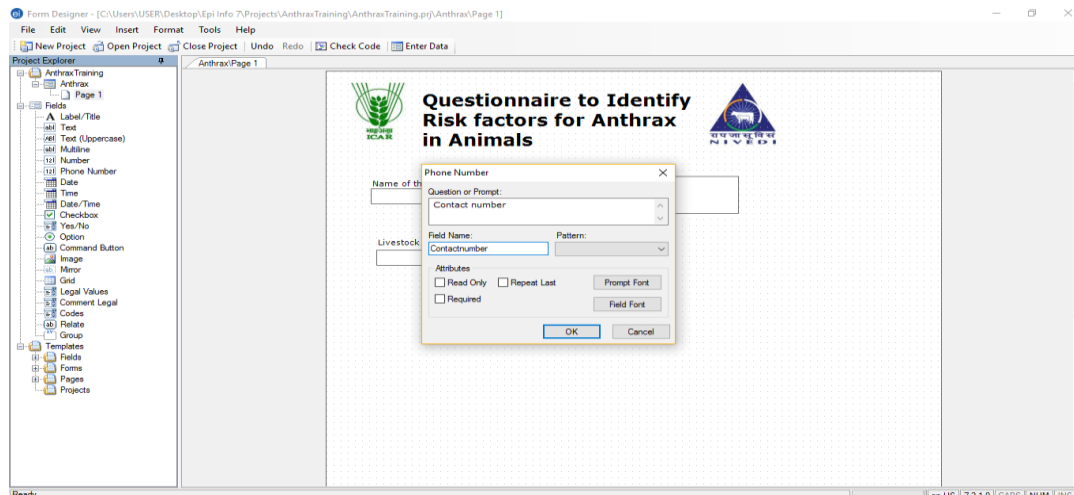


Figure 10

Yes/No: If your expected answer is either Yes or No then right click on the open space and click on Yes/No option as shown in Figure 11 and type **Do you know about Anthrax?**

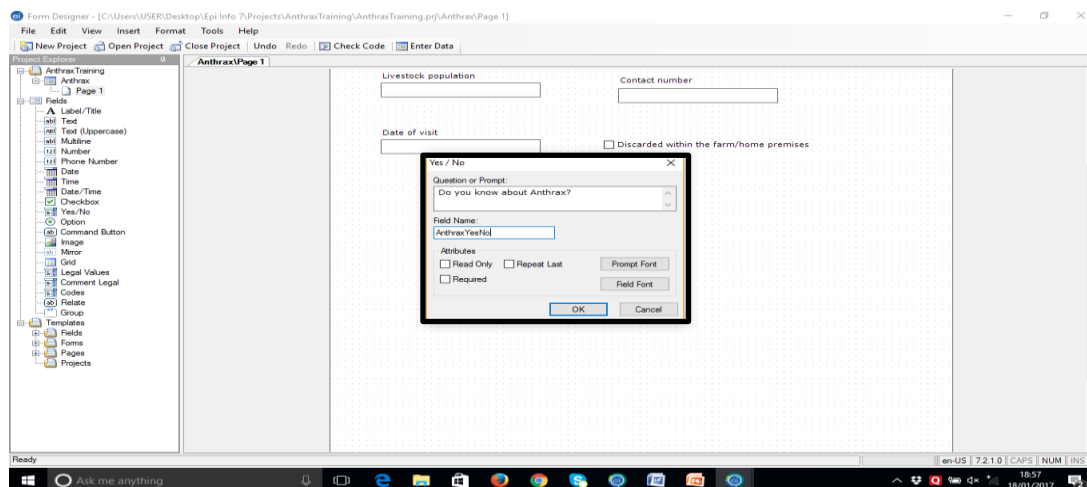


Figure 11

Option: If you want to have options with only one possibility then you can create an option variable (Figure 14). Then type **where you take your animals for grazing?**

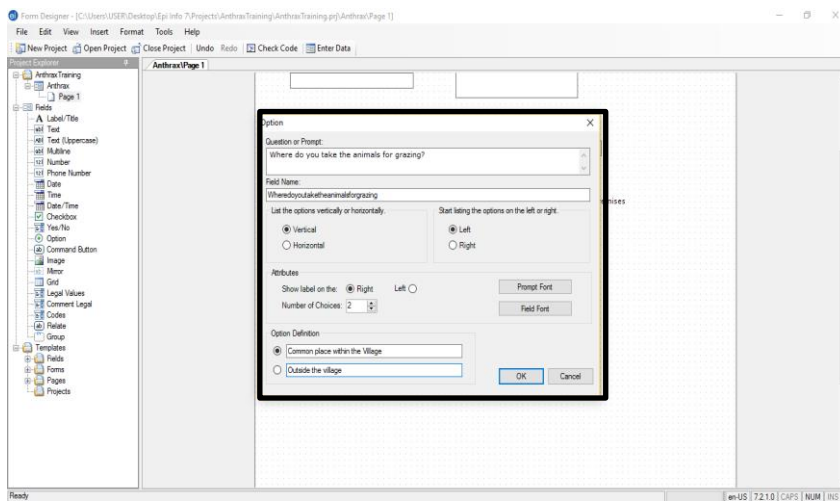


Figure 12

Legal Values: The option of legal values can be used when there are many options (e.g. Districts) and you want to create a drop down menu to select one of them. This can be done by right clicking and selecting **Legal Values** and type **Districts of Karnataka**. Give one field name (**Districts Karnataka**) and then click on Data source option (Figure 13). Then you will have option to add district names (Figure 14).

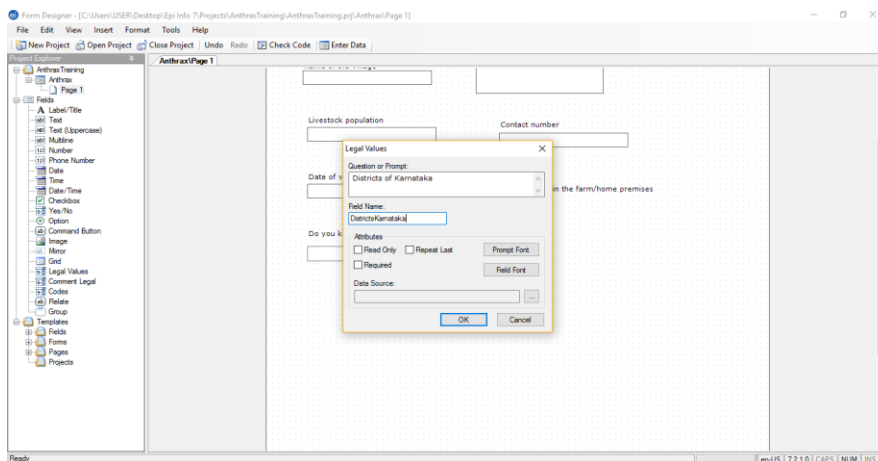


Figure 13

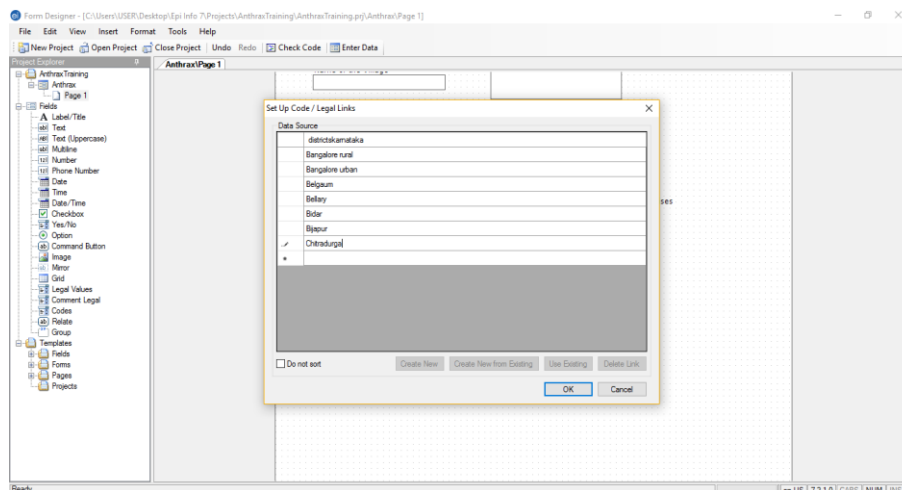


Figure 14

Once you have created legal values, you can try adding more fields to your questionnaire by right clicking on the database view of the **FORM**. If there is no space on the page then do not add more fields to your questionnaire. If you want to add more fields then you can add page by clicking on the option **Insert** (Figure 15).

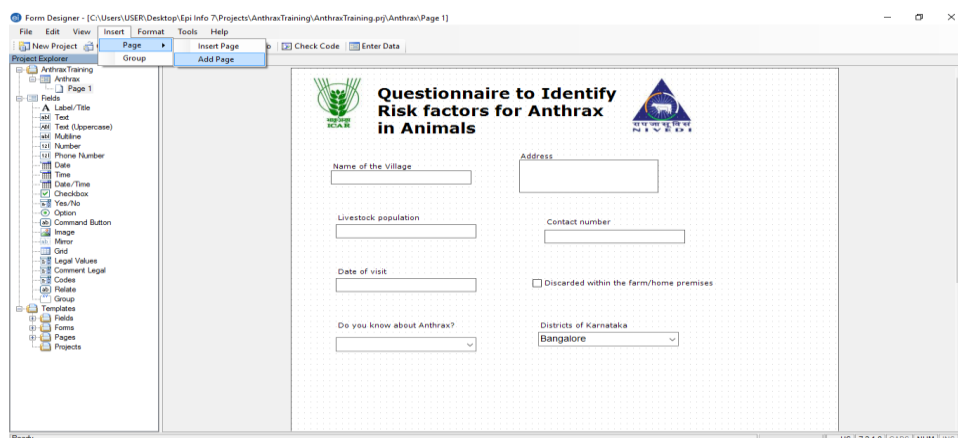


Figure 15

Tab order: Finally once your questionnaire is completed you have to check your tab order which will be used while entering data. Again right click and select **Tabs** and select on **Show Tab Order** (Figure 16). This will give you the tab order in which the questions are arranged (Figure 17).

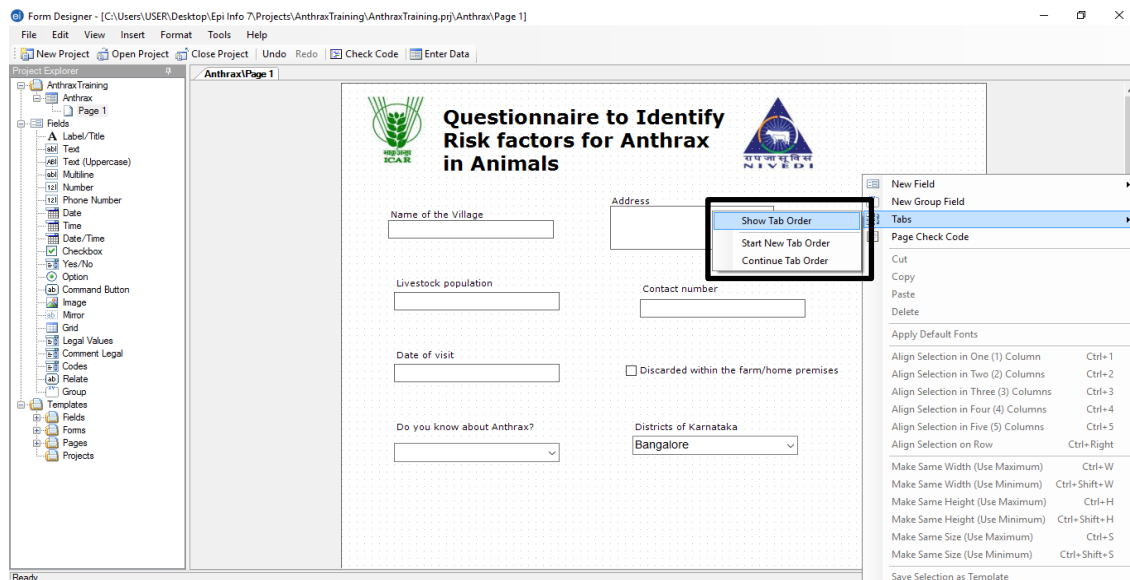


Figure 18

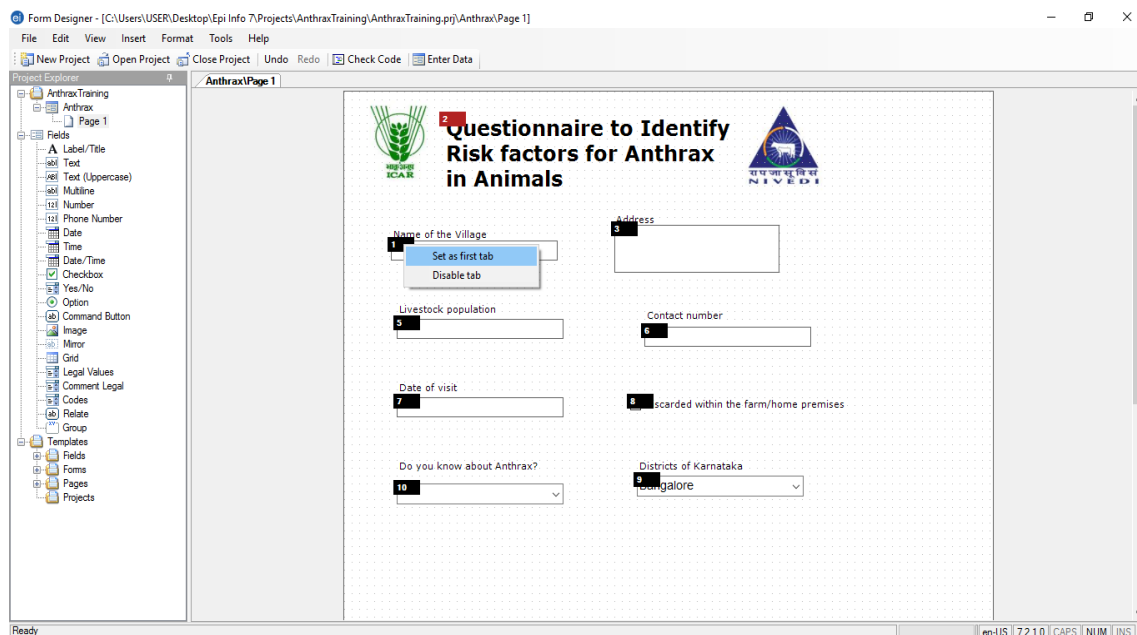


Figure 19

3. Entering of Field Collected Questionnaire Data in Epi-Info

Once you have collected the questionnaire data from field, it has to be entered in *Epi-Info* for further analysis.

Requirement for this practical

1. You should have *Epi-Info* questionnaire database created as discussed in previous chapter.
2. You should have hard copies of the questionnaire data collected from field
3. You should have created a questionnaire database in Epi-Info and same should be used for entering the data

Learning outcome of this Practical

1. You should be able to enter data in *Epi-info* software
2. You should be able to delete any in-correct entry
3. You should be able to create new record and enter data
4. You should be able to edit your entered data

Entering of data



Once you click on the **Epi-info** icon you will have options on the main dashboard. You click on the **ENTER DATA** option

Open form: Once you click on the **ENTER DATA** icon of the epi-Info software you will get a blank screen. You have to click on **Open Form** option (Figure 1) and open your **Form** created **Anthrax Training** and click on the form **Anthrax** (Figure 2).

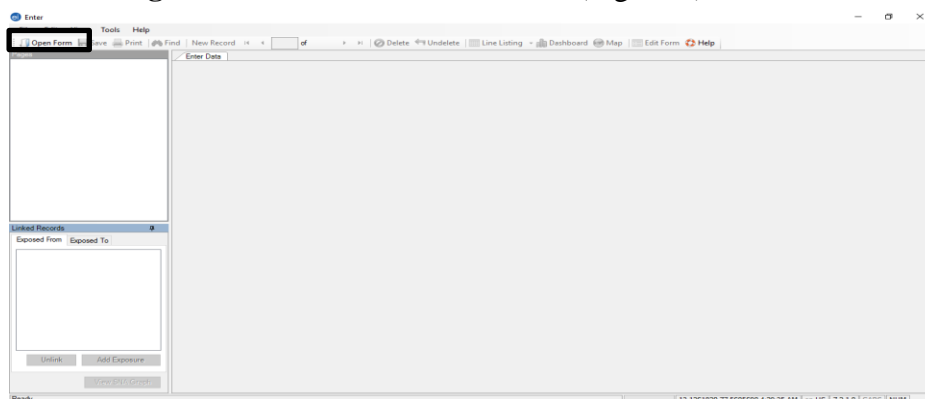


Figure 1

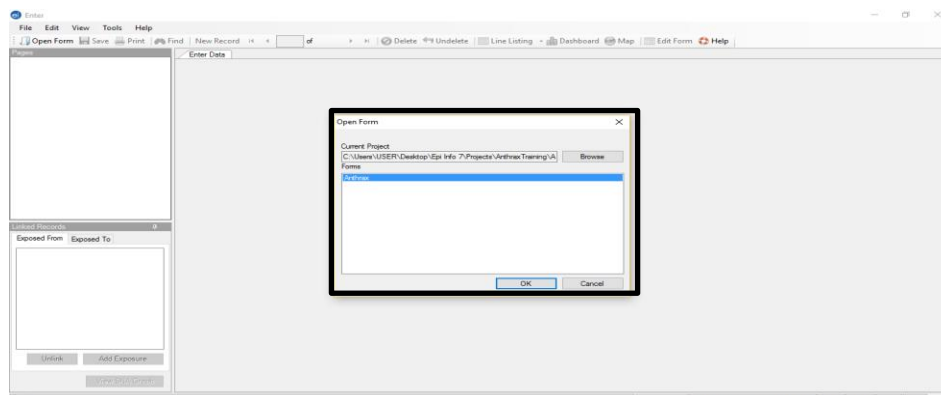


Figure 2

Once you click on the form Anthrax you will be asked to give **Data Table Name** in **Data Table Properties** and **Starting ID** for your records to be entered (Figure 3). This means that the first questionnaire will be given the Starting ID as 1 and subsequently when new records are added it will give numbers. Click on **OK.**]

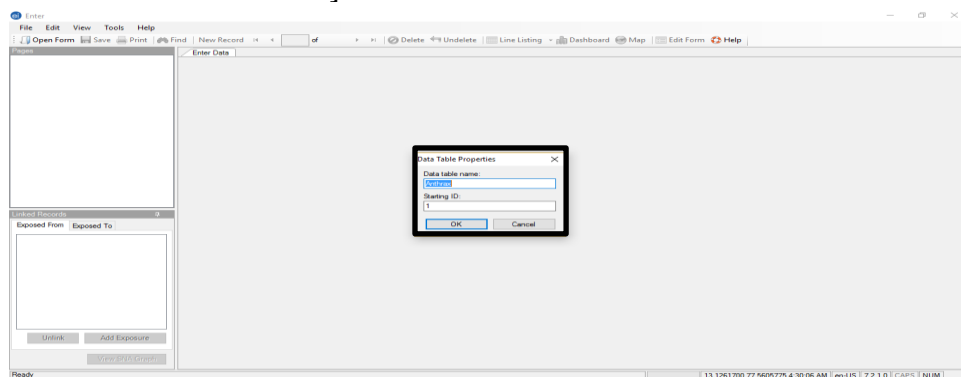


Figure 3

Entering data: Once you open the form you will get the questionnaire you created in **FORM** view and now you can enter the data (Figure 4). Start entering the data by giving your name and address or any other dummy information (Figure 5). Remember to **save** every time you enter the data.

The screenshot shows the 'Enter - [AnthraxTraining\Anthrax]' window. The main area displays a questionnaire titled 'Questionnaire to Identify Risk factors for Anthrax in Animals'. The form is empty, with fields for 'Name of the Village', 'Address', 'Livestock population', 'Contact number', 'Date of visit' (DD/MM/YYYY), and a checkbox for 'Discarded within the farm/home premises'. The sidebar on the left shows 'Page 1' and 'Linked Records'. The status bar at the bottom indicates '13.1261700.77.5605775.4.30.06 AM | en-US | 7.2.1.0 | CAPS | NUM'.

Figure 4

The screenshot shows the same questionnaire form as Figure 4, but with sample data entered. The 'Name of the Village' field contains 'Vijaykumar', 'Address' contains 'Kodihalli, doddaballapur, Bengaluru Rural, Karnataka', 'Livestock population' contains '2000', 'Contact number' contains '9999999999', 'Date of visit' contains '03/02/2017', and 'Do you know about Anthrax?' contains 'Yes'. The 'Discarded within the farm/home premises' checkbox is checked. The status bar at the bottom indicates '13.1262137.77.5605720.4.33.39 AM | en-US | 7.2.1.0 | CAPS | NUM'.

Figure 5

New record: Once you have entered the details of one farmer/respondent then you have to click on **New Record** (Figure 6). Once you click on new record a blank database will be seen on your screen (Figure 7). Here you can enter the details of the second farmer/respondent. Every time you have new respondent you need to click on New Record to enter the data

Enter - [AnthraxTraining\Anthrax]


File Edit View Tools Help

Open Form Save Print Find **New Record** of 0 Delete Undo Line Listing Dashboard Map Edit Form Help


Pages

- Anthrax
 - Page 1

Page 1



Questionnaire to Identify Risk factors for Anthrax in Animals



Name of the Village
Vijaykumar

Address
Kodihalli, Doddaballapur, Bengaluru Rural, Karnataka

Livestock population
2000

Contact number
9999999999

Date of visit
03/02/2017

☒ Discarded within the farm/home premises

Do you know about Anthrax?
Yes

Districts of Karnataka
Bangalore

Unlink Add Exposure View SNA Graph

[Name: DistrictsKarnataka | Type: LocalValues]

13.1262400 77.5605600 4:34:10 AM en-US 7.2.1.0 CAPS NIIM

Figure 6

Enter - [AnthraxTraining\Anthrax]


File Edit View Tools Help

Open Form Save Print Find New Record of 1 Delete Undo Line Listing Dashboard Map Edit Form Help


Pages

- Anthrax
 - Page 1

Page 1



Questionnaire to Identify Risk factors for Anthrax in Animals



Name of the Village

Address

Livestock population

Contact number

Date of visit
DD/MM/YYYY

☐ Discarded within the farm/home premises

Do you know about Anthrax?

Districts of Karnataka

Unlink Add Exposure View SNA Graph

[Name: Village | Type: Text]

New Record en-US 7.2.1.0 CAPS NIIM

Figure 7

4. Analysis of Epidemiological Data Using Visual Dashboard

Requirement for this practical

1. You should have data in the correct format to be analyzed in *Epi-info*
2. The data can be in excel, access or *Epi-info* format
3. It is not compulsory that the database should be created only in *Epi-Info* for analysis using visual dashboard. If the data is already entered in excel that can also be used for analysis
4. The objective and purpose of the analysis should be clearly defined. We can generate descriptive information (e.g. mean, frequency, pie charts, bar diagrams etc) or identify risk factors using Odds ratio or both.

Learning outcome of this Practical

1. You should be able to import data from excel or other formats into epi-info visual dashboard
2. You should be able to carry out descriptive analysis of your data e.g. calculate means, frequency, bar charts etc
3. You should be able to calculate Odds ratio/relative risk to identify potential risk factor for a disease

Analysis of data using visual Dashboard



Once you click on the **epi-info** icon you will have options on the main dashboard. You click on the **VISUAL DASHBOARD** option

Options in Visual Dashboard: Once you click on the Visual Dashboard icon you will get the Dashboard screen (Figure 1) with options as explained in the figure. The first step in analysis using Visual Dashboard is to **set data source**. There are three ways you can **set data source**. First using the menu, second using **set a data source now** and third by right clicking on the open space (Figure 2).

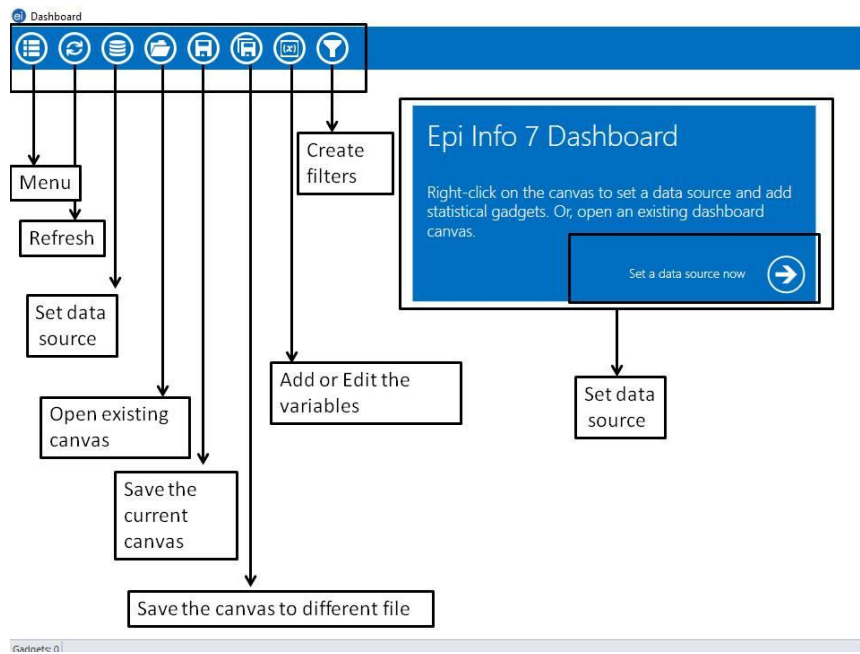


Figure 1

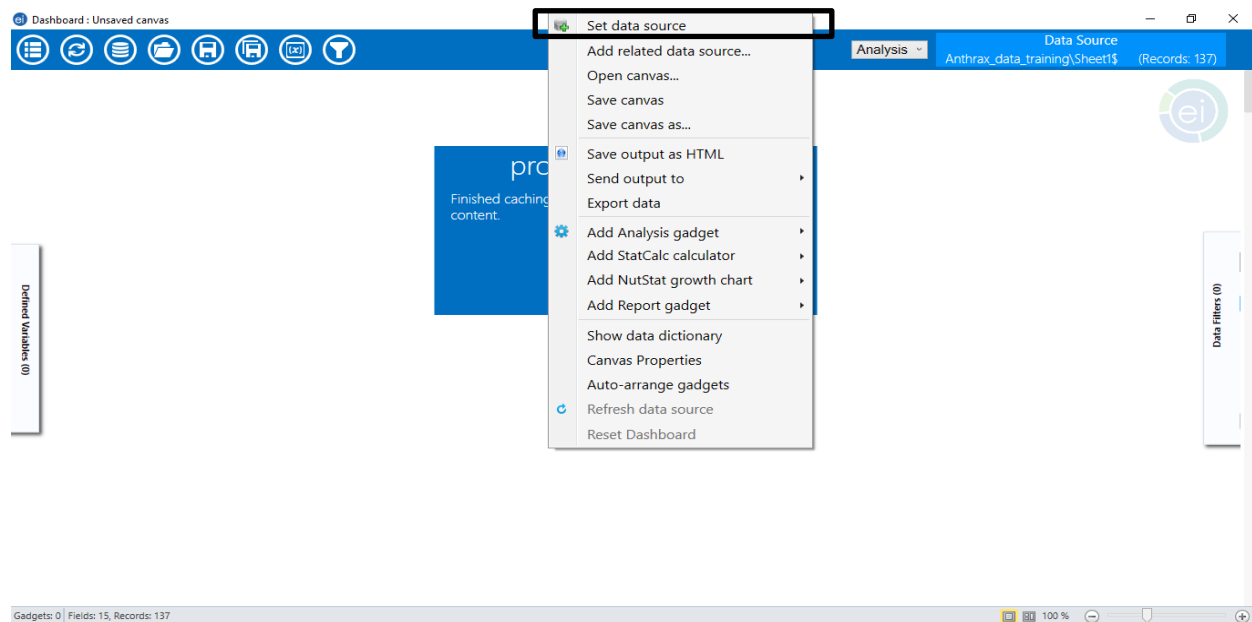
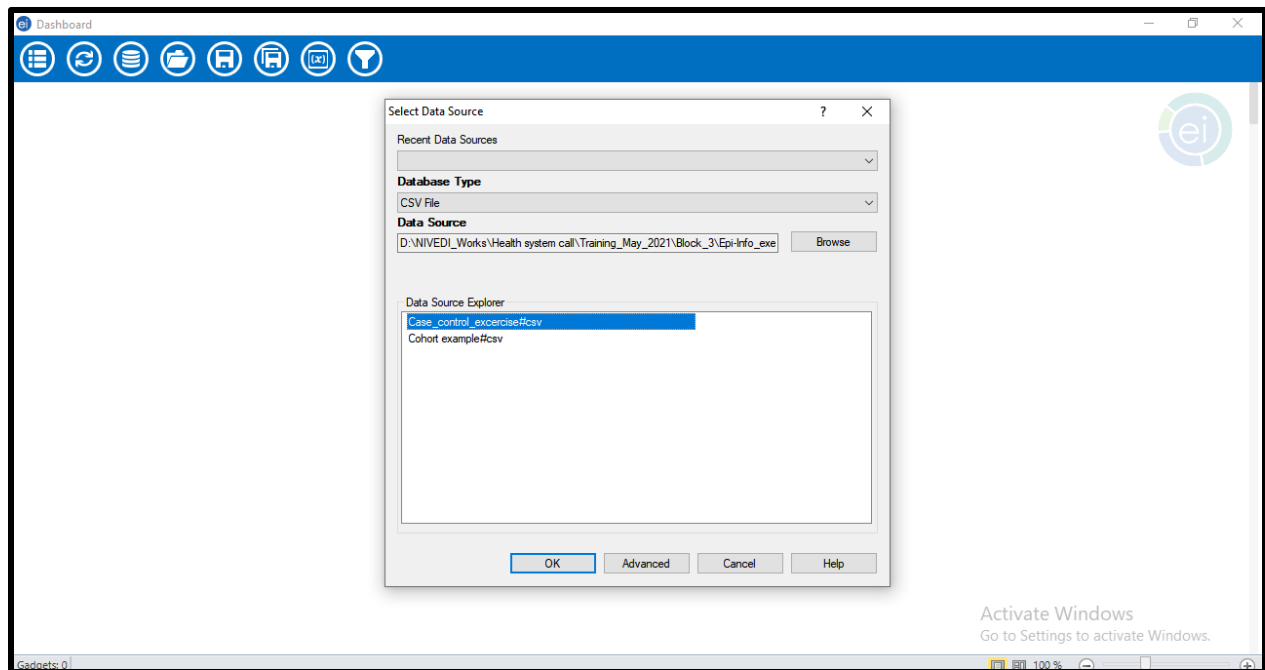


Figure 2



Once you click on the set data source option you will be asked to give the path where your data is stored. Here the file is named as **Case_control_exercise**. Select this data source and you can start your analysis.

Analysis of data: Once you right click on the dashboard you will get option of **Add Analysis gadget** (Figure 3). In this you will have options to do analysis depending on your objective of the analysis and kind of data source you have.

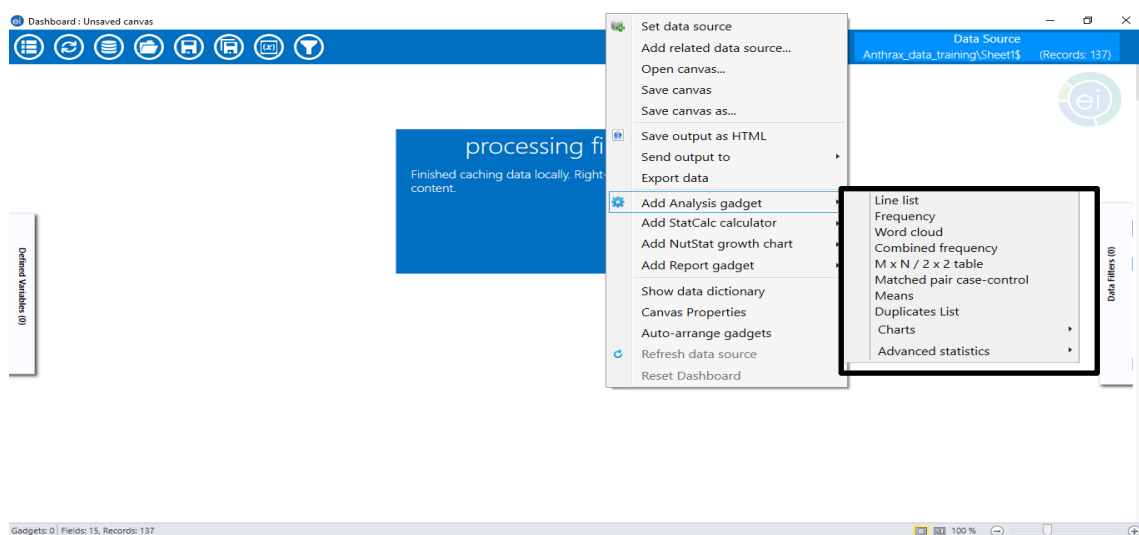


Figure 3

Frequency: If you to know the frequency of your variable under study then you can select **Frequency** from your **Add Analysis gadget** as shown in the figure 4.

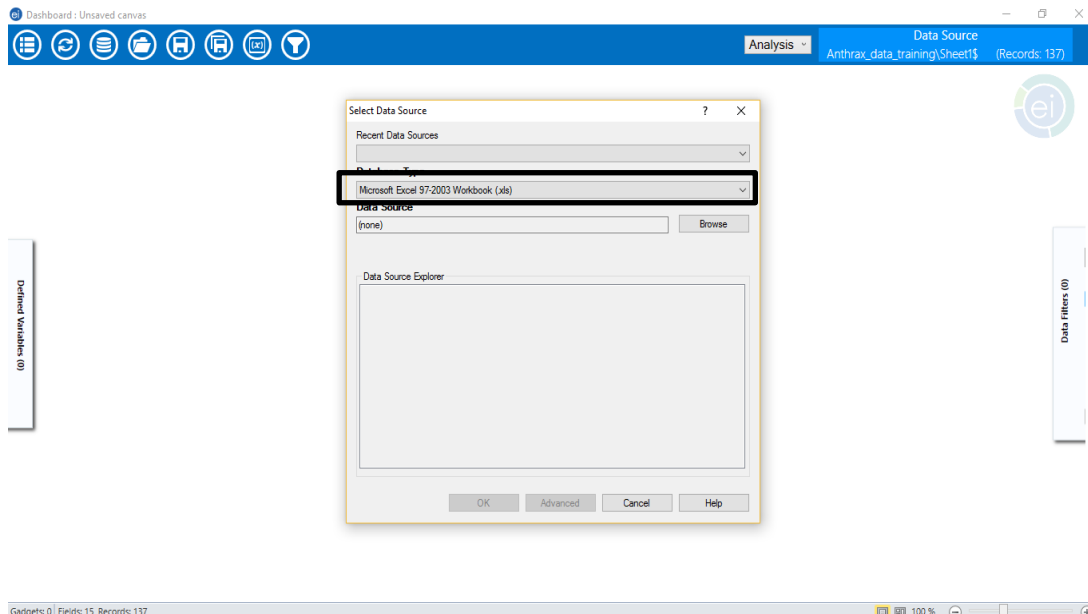


Figure 4

Once you click on the **Frequency** option you will get a screen to select the variables and also the options of grouping your data, filtering and display options (Figure 5). Select the variable **Death reported on neighbouring pasture** and click on OK.

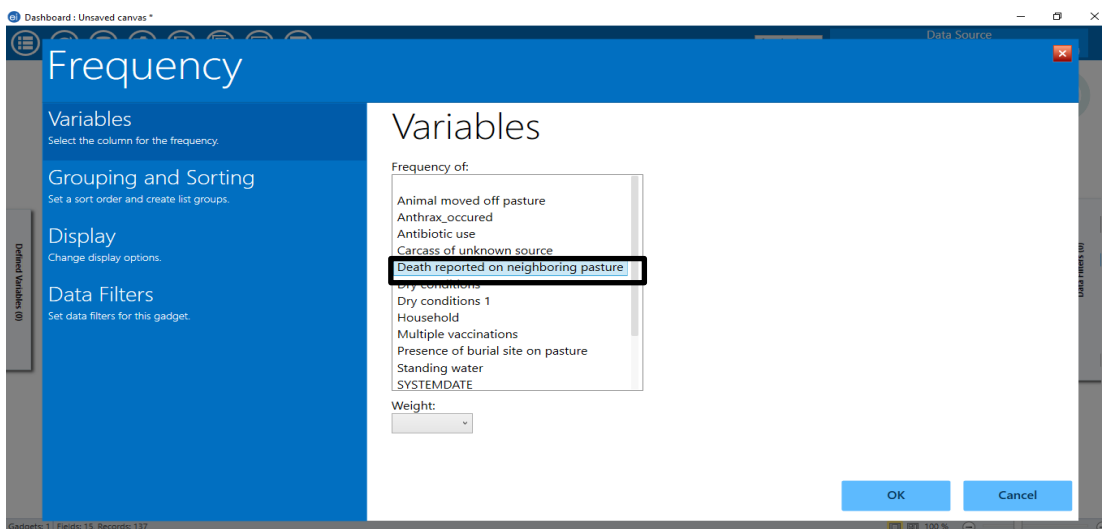



Figure 5

You will get the frequency of the **Death reported on neighbouring pasture** (Figure 6). If you want to make any changes to this frequency (e.g. Display) you can click on the  of the analysis and make changes.

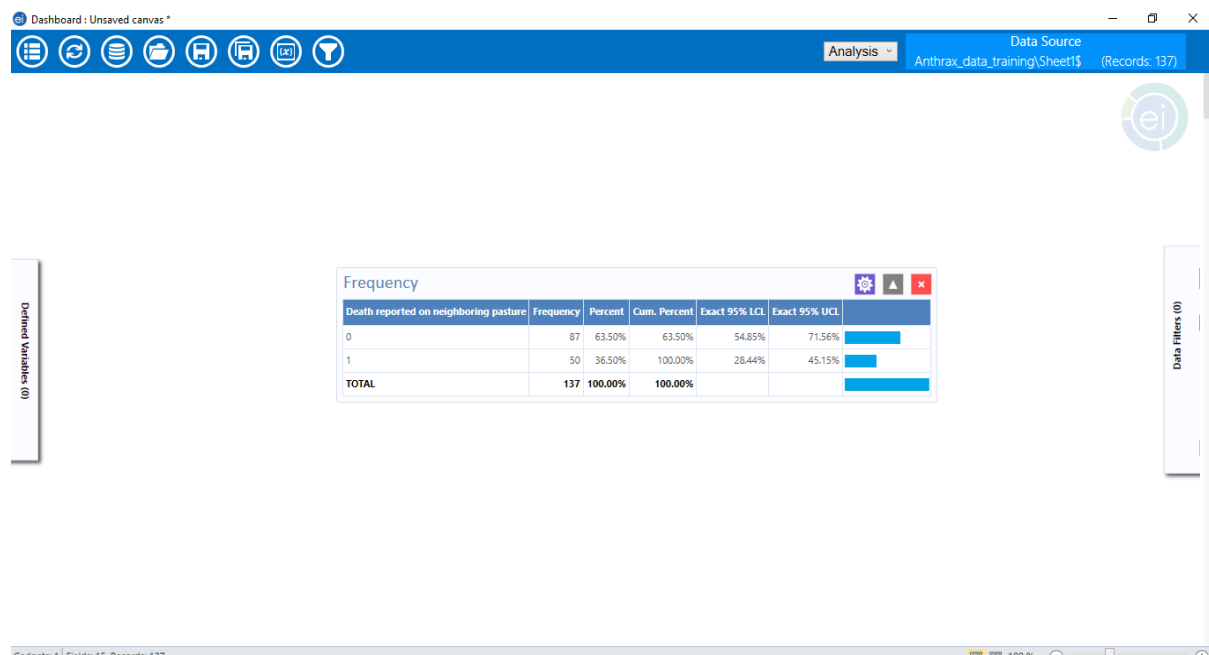


Figure 6

Odd's ratio: If you want to calculate Odd's ratio then right click select **MxN/2x2 table** from your **Add Analysis gadget** as shown in the figure 7.

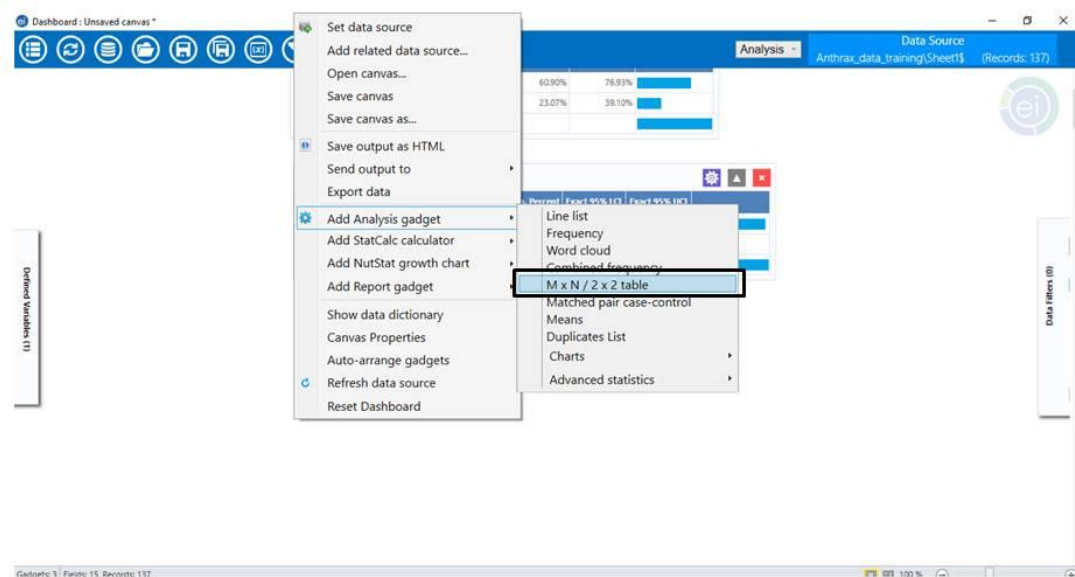


Figure 7

Once you select **MxN/2x2 Table** you have to select options of your variables (Figure 8). Select **Presence of burial site on pasture** as your Exposure variable and **Anthrax_occured** as your outcome variable and click on **OK**.

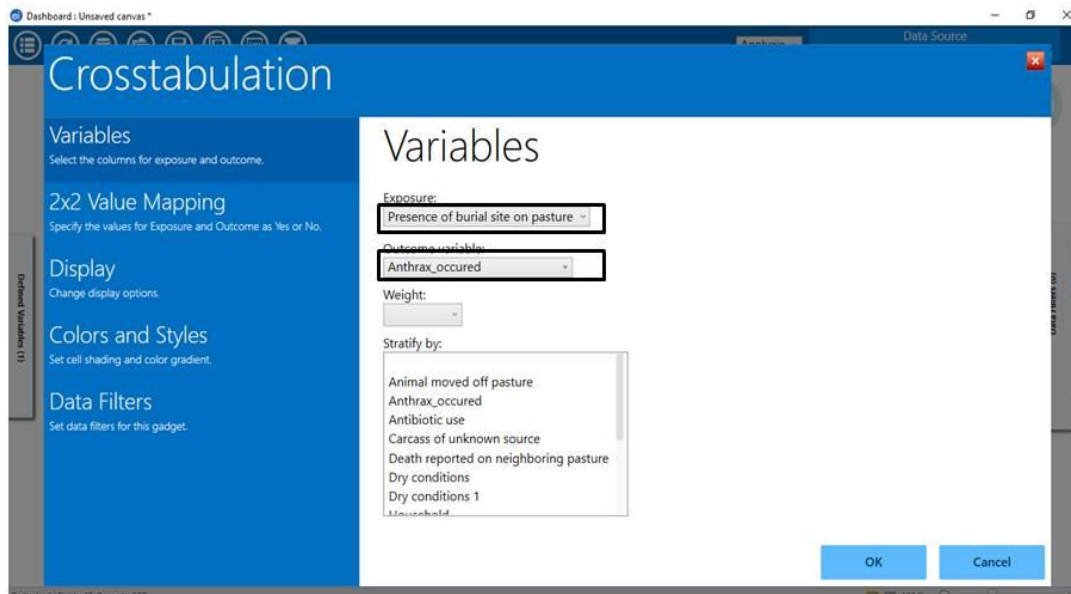


Figure 8

Once you click on OK you will get the results as shown in Figure 9. You can calculate Odd's ratio for other variables by repeating the above steps.

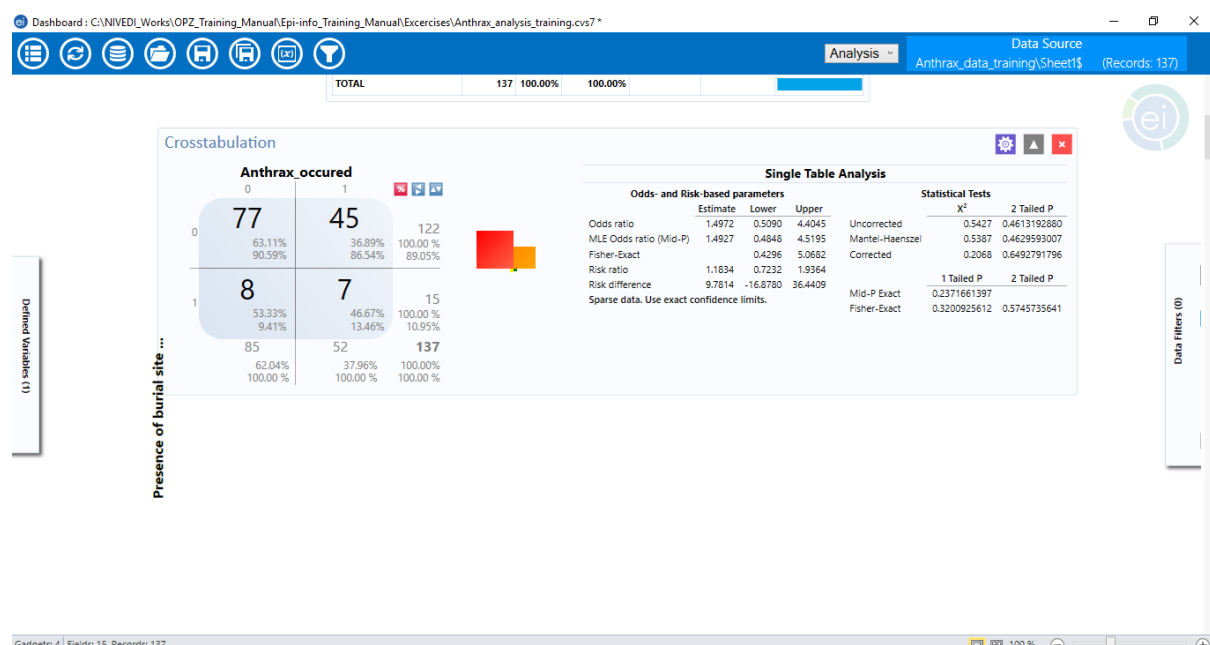


Figure 9

5. Exercises

Exercise 1: Identify risk factors for occurrence of Anthrax using case-control study design

The case study was developed by using hypothetical values for the risk factors and hence is not real risk factors. However, the probable risk factors and to create hypothetical data we referred this paper (Mongoh, M. N., Dyer, N. W., Stoltenow, C. L., & Khaita, M. L. (2008). Risk factors associated with anthrax outbreak in animals in North Dakota, 2005: A retrospective case-control study. *Public Health Reports*, 123(3), 352-359.)

Case study: Anthrax outbreak occurred from July 1 to October 12 2005 in 109 farms of 16 counties. The outbreaks were confirmed by Laboratory diagnosis (staining, isolation and PCR). Questionnaire was developed and sent to 419 farms, which included 109 farms confirmed for Anthrax and 130 farms negative for Anthrax. Out of 419 questionnaires 137 responded, which included 52 confirmed farms (Cases) and 85 negative farms (Control). Species affected were Bison, Cattle, Horses, Elk, Sheep, Deer.

Risk factors: 23 risk factors were considered for the questionnaire preparation and out of which 10 were selected as important by selection method. The hypothetical data on 10 variables are given to calculate Odd's ratio and your interpretation.

Study design: Retrospective case control study

Total respondents: 137

Cases: 52

Controls: 85

Exercise: Farms which were exposed to the risk factor and not exposed to the risk factor is given in the excel file. The data provided is hypothetically created for this exercise.

- a) Calculate the Odd's ratio for the risk factor
- b) Interpret the Odd's ratio

2. A retrospective cohort study of an outbreak of cryptosporidiosis among veterinary students

The case study was developed by using hypothetical values for the risk factors and meant for practicing in this exercise. We used this paper for creating hypothetical raw data; **Benschop, J., Booker, C. M., Shadbolt, T., & Weston, J. F. (2017). A Retrospective Cohort Study of an Outbreak of Cryptosporidiosis among Veterinary Students. *Veterinary sciences*, 4(2), 29.**

Objective of the study: Investigation of an outbreak of cryptosporidiosis among veterinary science and veterinary technology student at Massey University over an eight-week period during the spring-calving period in 2013

Case definition: a clinically-confirmed case was defined as a veterinary science or veterinary technology student from Massey University with diarrhoea (of at least 24 hours duration) and/or vomiting where the onset of illness was between 20 August 2013 and 10 October 2013.

A laboratory-confirmed case was a clinically-confirmed case that tested positive for *Cryptosporidium*

Questionnaires

Questionnaires were distributed to second- and third-year students (n=56) and final year students (n=100)

Exercise

Perform descriptive analyses of all the probable risk factors

Calculate the relative risk for all the risk factors