Case Study 12: Swaziland Farm Animal Genetic Resources Survey: estimation of population numbers for different livestock species

Moses Mefika Sithole\(^1\), Dorah Vilakati\(^2\) and John Rowlands\(^3\)

\(^1\)Department of Agricultural Economics and Management, University of Swaziland, P.O. Luyengo, Swaziland
\(^2\)Ministry of Agriculture and Cooperatives, Swaziland Government, P.O. Box 162 Mbabane, Swaziland
\(^3\)International Livestock Research Institute, P.O. Box 30709, Nairobi, Kenya

Contents

Summary
Background
Study design
Objectives
Questions to be addressed
Source material
Exploration & description
Findings, implications and lessons learned
Study questions
Related reading
Acknowledgements

Summary

This case study uses data on the numbers of livestock species in homesteads, collected in a livestock breed survey implemented in Swaziland, to estimate population numbers and their standard errors. Whilst the survey covered the whole of Swaziland, the results reported here are just for the Hhohho region used in Case Study 11.

Sampling was done in two stages. Firstly, dip-tank areas were selected for sampling in each sub-region, taking into account such factors as livestock density, proximity to commercial farms and ecological zone; secondly, homesteads were randomly sampled within each sampled dip-tank area.

This case study shows how estimates of the numbers of livestock in the Hhohho region (and their standard errors) can be calculated. This has been done by using different formulae for calculating standard errors, assuming that the selection of dip-tank areas was random.

The two contrasting situations, whereby the numbers of homesteads in each dip-tank area throughout the region (whether sampled or not) are known or not known, are also analysed. This is followed by a discussion on the effect of different methods of sampling, not only on the methods of calculation but also on the reliability of the population estimates that are derived.

Chickens, followed by cattle and goats, in that order, were kept in the largest numbers by
smallholder farmers.

**Background**

This case study is based on a livestock breed survey implemented in Swaziland and described in Case Study 11, which described approaches to data analysis in relation to the management of and purposes for keeping cattle by smallholder farmers.

As described in that case study, a lack of knowledge of the physical and performance characteristics of indigenous farm animal genetic resources in countries in sub-Saharan Africa, and the extent of existing genetic diversity has, through pressures to increase production, led to the underutilisation, dilution and replacement of these resources through crossbreeding with exotic breeds.

In order to obtain better knowledge of existing indigenous breeds of cattle, sheep, goats, chickens, pigs and donkeys, and the management thereof, FAO, together with ILRI, planned a series of surveys to gather this information across the Southern African Development Community (SADC) region, including Swaziland.

One objective was to obtain estimates of the total numbers of livestock in the country. This case study describes methods for obtaining such estimates, when applied to one of the regions in Swaziland, and provides some preliminary estimates of totals and confidence ranges for the six livestock species.

The sampling framework is outlined in Case Study 11; however, certain key features are repeated here.

The country is divided into four administrative regions - Hhohho, Manzini, Shiselweni and Lubombo (primary strata). These regions represent distinct socio-economic patterns and modes of livelihood. Within each of these regions there are 28 livestock sub-regions (secondary strata) - 6, 7, 7 and 8 for the four regions, respectively. It had been decided to include each of the 28 sub-regions in the survey and select dip-tank areas from each sub-region.

The homesteads that kept livestock in Swaziland (approximately 180,000) were defined as the population for the survey. The numbers of homesteads in each dip-tank area were known from census data obtained before the commencement of the survey.
Study design

Sampling of homesteads was done in two stages of cluster sampling.

In the first stage, a mixture of representative and random sampling of dip-tank areas (primary sampling units - PSUs) was done within each sub-region. The representative nature of the sampling took account of ecological zone, proximity to commercial farms, proximity to urban areas, remoteness and livestock density. Remoteness was taken into account in order not to miss a genetic resource that might be present only in one particular location.

Whilst representative sampling has the attraction that it ensures that all situations can be covered, it has the disadvantage that a standard error cannot be calculated for the estimate of the population total. Furthermore, the estimate of the population total may be biased.

One way round the problem is to define strata within each sub-region based on the criteria mentioned above, and then to sample dip-tanks areas at random from within each stratum. This method was used in this survey.

The numbers of dip-tank areas selected within a sub-region depended on the total number of dip-tank areas and presumed livestock density in that sub-region; in general, the more dip-tank areas the larger was the sample. Thus, there was an attempt to sample from sub-regions approximately, but not entirely, proportional to the number of dip-tank areas.

For the purposes of this case study, and to ensure that formulae do not become too complicated, we shall ignore stratification. We shall also assume two scenarios during the course of this case study, firstly that dip-tank areas are selected at random with equal probability, and, secondly that they are selected with probability proportional to their size.

Dip-tank areas were used as clusters in the sampling frame because they form an important feature within the livestock administrative structure for Swaziland. Indeed, the majority of the government’s livestock departmental field staff is based in dip-tank areas.

Taking into consideration cost, human resource and time, it was decided to sample a total of 60 dip-tank areas in each of the four regions, i.e. a total of 240 from the 430 dip tanks in the country.

The table shows the numbers of dip-tank areas within the sub-regions in the Hhohho region and the numbers of dip-tank areas sampled.

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Population</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lobamba</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Malandzea</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Mayiwane</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Mbabane</td>
<td>28</td>
<td>15</td>
</tr>
<tr>
<td>Nfonjeni</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Piggs Peak</td>
<td>14</td>
<td>5</td>
</tr>
</tbody>
</table>
In the second stage of the sampling procedure, homesteads (ultimate sampling units - USUs) were randomly sampled within selected dip-tank areas.

In order to adequately cover each dip-tank area, it was decided to sample as far as possible at least 10 homesteads per dip-tank area to give a total sample of 2,400 homesteads countrywide. This sample size (a sampling fraction of approximately 1.3%) is comparable to that normally used by the Central Statistical Office (CSO) in agricultural surveys.

The table shows the numbers of homesteads sampled within the sampled dip-tank areas of Sub-region 3 (Mayiwane) of the Hhohho Region.

<table>
<thead>
<tr>
<th>Dip-tank area</th>
<th>Homesteads with livestock</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Population</td>
</tr>
<tr>
<td>Bazara</td>
<td>197</td>
</tr>
<tr>
<td>Jacks</td>
<td>222</td>
</tr>
<tr>
<td>Lugongodiwa</td>
<td>141</td>
</tr>
<tr>
<td>Mabuko</td>
<td>140</td>
</tr>
<tr>
<td>Magidza</td>
<td>198</td>
</tr>
<tr>
<td>Magonigoni</td>
<td>222</td>
</tr>
<tr>
<td>Mangololo</td>
<td>144</td>
</tr>
<tr>
<td>Manjengeni</td>
<td>171</td>
</tr>
<tr>
<td>Masheleleni</td>
<td>279</td>
</tr>
<tr>
<td>Mboma</td>
<td>180</td>
</tr>
<tr>
<td>Ndalambi</td>
<td>238</td>
</tr>
<tr>
<td>Ntsinini</td>
<td>214</td>
</tr>
<tr>
<td>Shumi</td>
<td>116</td>
</tr>
<tr>
<td>Sidwashini</td>
<td>249</td>
</tr>
<tr>
<td>Zinyane</td>
<td>98</td>
</tr>
<tr>
<td>Total</td>
<td>2809</td>
</tr>
<tr>
<td>Hhohho Region</td>
<td>9047</td>
</tr>
</tbody>
</table>

Some questionnaires had to be discarded due to incomplete or missing information.

Sometimes prior information on the different species kept by a particular homestead was found to be incorrect see (Case Study 11). Thus, on reaching a homestead, supposedly keeping pigs and pre-selected as the primary species to be completed for the questionnaire, the enumerator may have found that the homestead kept other species, but not pigs.

More than 10 homesteads were interviewed in some dip-tank areas. Use of a modified random walk method for selection of homesteads sometimes led to additional homesteads being interviewed.

Objectives

This case study has three primary objectives:

- To provide preliminary estimates of the numbers of cattle, sheep, goats, pigs, chickens and donkeys in Hhohho Region of Swaziland.
- To compare alternative formulae for estimating population numbers and calculating standard errors.
- To discuss different methods of survey design that will help to optimise reliability and precision of population estimates.
Questions to be addressed

The following topics in survey design will be addressed:

- Methods of deriving population estimates and their standard errors.
- The importance of understanding the way that a sampling frame has been planned when estimating population values.
- The importance of ensuring as far as possible random sample selection.
- The value of stratification.
- Possible use of ancillary information in improving precision of population estimates.

We use Microsoft Excel to demonstrate the calculations of population totals and confidence ranges.

Source material

The complete data set used for the case study is held in Breedsurv - a computer data capture and storage system designed and written in Microsoft Access 2000 by ILRI specifically for SADC Livestock Breed Surveys (Rowlands et al., 2003). This is briefly described in Case Study 11.

Only the data for the numbers of the different species recorded on Page 1c of the questionnaire (see Case Study 11) are required for this case study. These are stored in a Breedsurv table called HOUSEHOLD.
The livestock numbers have been extracted and copied into duplicate worksheets stored in the Excel 'Data' worksheets in CS12Data1 and CS12Data2. The different field names in these data worksheets are described in CS12Doc.

These two data sets are used for different methods of calculating population estimates.

Source: Thalile Seipone Sowane

Exploration and description

Before giving the formulae for calculating population estimates, it is helpful to see again the details of the numbers of dip-tank areas in the different sub-regions, and the numbers of homesteads within dip-tank areas that possessed livestock. Firstly, here is the table that gives the numbers of dip-tanks in the sub-regions and the numbers sampled.

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Population (N)</th>
<th>Sample (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lobamba</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Malandzela</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Mayiwane</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Mbabane</td>
<td>28</td>
<td>15</td>
</tr>
<tr>
<td>Ntonjeni</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Pigg's Peak</td>
<td>14</td>
<td>5</td>
</tr>
</tbody>
</table>

Population and sample sizes

And here on this and the next two pages are details of the numbers of homesteads with livestock in the six sub-regions.

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Homesteads in sampled dip-tank areas</th>
<th>Homesteads in non-sampled areas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dip-tank area</td>
<td>Population</td>
</tr>
<tr>
<td>Lobamba</td>
<td>Buhlungu</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Ezulwini</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>Luphahlane</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>Mhlambanyatsi</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
And here on this and the next two pages are details of the numbers of homesteads with livestock in the six sub-regions.

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Homesteads in sampled dip-tank areas</th>
<th>Homesteads in non-sampled areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lobamba</td>
<td>Dip-tank area</td>
<td>Population</td>
</tr>
<tr>
<td></td>
<td>Buhlungu</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Ezulwini</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>Luphahlane</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>Mhlambanyatsi</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Malandzela</th>
<th>Homesteads in sampled dip-tank areas</th>
<th>Homesteads in non-sampled areas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dip-tank area</td>
<td>Population</td>
</tr>
<tr>
<td></td>
<td>Esifutfweni</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Logobisa</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Malandzela</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>Maphalaleni</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>Melethi</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>Ntoma</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>Steyn's</td>
<td>163</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Homesteads in sampled dip-tank areas</th>
<th>Homesteads in non-sampled areas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dip-tank area</td>
<td>Population</td>
</tr>
<tr>
<td></td>
<td>Bazara</td>
<td>197</td>
</tr>
<tr>
<td></td>
<td>Jacks</td>
<td>222</td>
</tr>
<tr>
<td></td>
<td>Lugongodwa</td>
<td>141</td>
</tr>
<tr>
<td></td>
<td>Maboko</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>Magidza</td>
<td>198</td>
</tr>
<tr>
<td></td>
<td>Magonigoni</td>
<td>222</td>
</tr>
<tr>
<td></td>
<td>Mangololo</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td>Manjengeni</td>
<td>171</td>
</tr>
<tr>
<td></td>
<td>Masheleleni</td>
<td>279</td>
</tr>
<tr>
<td></td>
<td>Mboma</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>Ndlalambi</td>
<td>238</td>
</tr>
<tr>
<td></td>
<td>Ntsinini</td>
<td>214</td>
</tr>
<tr>
<td></td>
<td>Shumi</td>
<td>116</td>
</tr>
<tr>
<td></td>
<td>Sidwashini</td>
<td>249</td>
</tr>
<tr>
<td></td>
<td>Zinyane</td>
<td>98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ntfonjeni</th>
<th>Homesteads in sampled dip-tank areas</th>
<th>Homesteads in non-sampled areas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dip-tank area</td>
<td>Population</td>
</tr>
<tr>
<td></td>
<td>Elahleni</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>Mawasha</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Mawelawela</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>Mthondwane</td>
<td>335</td>
</tr>
<tr>
<td></td>
<td>Nkonjaneni</td>
<td>232</td>
</tr>
<tr>
<td></td>
<td>Ntintinyane</td>
<td>275</td>
</tr>
<tr>
<td></td>
<td>Vuseni</td>
<td>171</td>
</tr>
</tbody>
</table>
Let us start by showing how to estimate the total number of cattle in a single dip-tank area. We shall base our calculations on formulae given by Rowlands et al. (2003) for a similar survey in Zimbabwe.

Let us consider dip-tank area \( i \), where \( i \) represents one of the dip-tank areas sampled. Suppose that \( m_i \) homesteads are sampled in this dip-tank area and that there are \( M_i \) homesteads in total in this area. Suppose that \( y_{ij} \) animals are recorded per homestead (\( j = 1 \) ... \( m_i \)), and that \( \bar{y}_i \) represents the mean number of cattle per homestead. This can be calculated as:

\[
\bar{y}_i = \frac{\sum y_{ij}}{m_i}
\]

Then the estimate of the total number of cattle in the dip-tank area is

\[
T_i = M_i \bar{y}_i
\]

and the variance of \( T_i = M_i (M_i - m_i) s_i^2 / m_i \), where \( s_i^2 \) is the sample variance equal to

\[
\sum (y_{ij} - \bar{y}_i)^2 / (m_i - 1) \quad (j = 1 \ldots m_i).
\]

The standard error of the total in this dip-tank area is thus the square root of the variance

\[
\sqrt{M_i (M_i - m_i) s_i^2 / m_i}
\]

Thus, we can write the estimate of the total number of cattle in the dip-tank area as

\[
M_i \bar{y}_i \pm \sqrt{M_i (M_i - m_i) s_i^2 / m_i}
\]

Let us now extend these calculations to estimate the total number of cattle in the sub-region.
Let $T_1 \ldots T_n$ be the estimated totals of cattle in the $n$ dip-tank areas sampled, and $M_1 \ldots M_n$ be the numbers of homesteads in these dip-tank areas.

As we have already seen, the estimated total $T_i$ for the $i^{th}$ dip-tank area can be written: $T_i = \bar{Y}_i M_i (i = 1 \ldots n)$. Also, the variance of $T_i$ is given by $S_i^2 = M_i (M_i - m_i) s_i^2 / m_i$, where, as before, $m_i$ is the number of homesteads sampled in dip-tank area $i$, and $s_i^2$ is the variance of this sample.

An unbiased estimate of the average number of animals per homestead across the $n$ sampled dip-tank areas ($n=4$ in Lobamba - see three pages back) can then be written as $\bar{Y} = (T_1 + \ldots + T_n) / (M_1 + \ldots + M_n)$.

Suppose that there are $N$ dip-tank areas in the sub-region (N=13 in Lobamba) and that the number of homesteads with livestock $M_i$ is also known in each of the dip-tank areas that were not sampled. Then an unbiased estimate of the total for the sub-region, assuming that the sampled dip-tank areas were selected at random, is: $T = \bar{Y} (M_1 + \ldots + M_N)$.

However, if the numbers of homesteads in non-sampled dip-tank areas are not known, then the best unbiased estimate of the total is: $T = \bar{Y} (M_1 + \ldots + M_N) \times N/n$

The among dip-tank area within sub-region component of the variance of $T$ is calculated as $U = N(N-n) S_T^2$ where $S_T^2$ is the variance among $T_1 \ldots T_n$.

$S_T^2$ can be calculated in one of two ways:

(i) $S_T^2 = [(T_1 - T)^2 + \ldots + (T_n - T)^2] / (n-1)$ where $T = \sum (T_i)/n$

(ii) $S_T^2 = [(T_1 - \bar{Y} M_1)^2 + \ldots + (T_n - \bar{Y} M_n)^2] / (n-1)$

The choice of methods is described in the next page.

We also need to include the within dip-tank area variance components $S_1^2 \ldots S_n^2$ that have already been calculated.

Thus, the complete formula for the variance of $T$, the estimated total number of animals in the dip-tank area, is given by $\text{Var}(T) = U + N (S_1^2 + \ldots + S_n^2) / n$.

The standard error is the square root of this expression.

We have suggested two alternative formulae for $S_T^2$, namely the variance among the $T_i$s. We shall demonstrate their use in two different scenarios.

**Scenario 1**: dip-tank areas selected at random with equal probabilities.
\[ S_T^2 = \frac{(T_1 - T)^2 + \ldots + (T_n - T)^2}{n-1} \] where \( T = \frac{\sum T_i}{n} \)

This is an unbiased estimate but is likely to be comparatively large, especially if dip-tank areas vary in size. We shall also assume in this scenario that knowledge of the numbers of homesteads in a dip-tank area prior to the start of a survey is not available.

**Scenario 2**: dip-tank areas selected at random but proportional to dip-tank area size.

\[ S_T^2 = \frac{(\overline{Y_j} - \overline{M_j})^2 + \ldots + (\overline{Y_n} - \overline{M_n})^2}{n-1} \]

This estimate will tend to be smaller than in Scenario 1, especially if the average number of cattle per homestead \( \overline{Y_j} \) are relatively uniform across dip-tank areas. In contrast to Scenario 1 we shall assume that we know the numbers of homesteads in a dip-tank area for all dip-tank areas in a sub-region.

**Estimation**

**Scenario 1**

We show below the information (shaded in orange) that we need in order to apply the formulae given for Scenario 1 to Lobamba Sub-region.

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Dip-tank areas</th>
<th>Population (N)</th>
<th>Sample (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lobamba</td>
<td></td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Malandzela</td>
<td></td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Mayiwane</td>
<td></td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Mbabane</td>
<td></td>
<td>28</td>
<td>15</td>
</tr>
<tr>
<td>Ntfonjeni</td>
<td></td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Pigg's Peak</td>
<td></td>
<td>14</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Dip-tank area</th>
<th>Population</th>
<th>Sample</th>
<th>Dip-tank area</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lobamba</td>
<td>Buhlungu</td>
<td>105</td>
<td>5</td>
<td>Bhidilili</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td>Ezulwini</td>
<td>126</td>
<td>6</td>
<td>Mdonjane</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>Luphahlane</td>
<td>147</td>
<td>8</td>
<td>Sigcawini</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Mhlambanyatsi</td>
<td>126</td>
<td>7</td>
<td>Vezi</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Makhebelele</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Luyengweni</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mambiteni</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Magwama</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Majuba</td>
<td>60</td>
</tr>
</tbody>
</table>

First we can calculate \( \overline{Y_j} \) (mean) and \( s_j^2 \) (variance) for each dip-tank area.

This may be accomplished, for example for Lobamba Sub-region for cattle, by using the Pivot Table feature of Microsoft Excel for the variable cat_no in CS12Data1 (Data Pivot → Table and Pivot Chart Report...) as shown in `Descr_Stats' worksheet of CS12Data1 and
illustrated below.

Notice the ranges in numbers of cattle recorded per homestead. There is some evidence of skewness in the numbers of cattle (cat_no).

These results are then used to calculate an estimate of the size of the population of cattle in the four sampled sub-regions using the formula shown earlier:

\[ T_i = \sqrt{\bar{Y}_i M_i} \text{ and } \text{s.e.} \ (T_i) = \sqrt{S_i^2} = \sqrt{M_i (M_i - m_i) s_i^2 / m_i} \quad (i = 1 \ldots 4) \]

### Lobamba (with 4 of 13 dip-tank areas sampled)

<table>
<thead>
<tr>
<th></th>
<th>Buhlungu</th>
<th>Ezulwini</th>
<th>Luphahlane</th>
<th>Mhlambanyatsi</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{Y}_i )</td>
<td>22.4</td>
<td>19.3</td>
<td>6.9</td>
<td>13.6</td>
</tr>
<tr>
<td>s.e. of ( \bar{Y}_i )</td>
<td>8.8</td>
<td>6.9</td>
<td>1.9</td>
<td>2.1</td>
</tr>
<tr>
<td>( m_i )</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>( M_i )</td>
<td>105</td>
<td>126</td>
<td>147</td>
<td>126</td>
</tr>
<tr>
<td>( T_i ) (± s.e.)</td>
<td>2,352±904</td>
<td>2,436±851</td>
<td>1,011±269</td>
<td>1,710±253</td>
</tr>
</tbody>
</table>

The calculations are shown in the first part of the worksheet spreadsheet `Cat_no_pop_inf` of CS12Data1.

We can similarly calculate estimates for the other dip-tank areas.

For example, for Malandzela:

### Malandzela (with 7 of 12 dip-tank areas sampled)

<table>
<thead>
<tr>
<th></th>
<th>Esifutweni</th>
<th>Logobisa</th>
<th>Malandzela</th>
<th>Maphalaleni</th>
<th>Melethi</th>
<th>Ntima</th>
<th>Steyn's</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{Y}_i )</td>
<td>24.0</td>
<td>17.4</td>
<td>19.3</td>
<td>15.2</td>
<td>18.2</td>
<td>15.2</td>
<td>7.7</td>
</tr>
<tr>
<td>s.e. of ( \bar{Y}_i )</td>
<td>9.2</td>
<td>3.7</td>
<td>7.3</td>
<td>4.1</td>
<td>3.6</td>
<td>2.4</td>
<td>11.4</td>
</tr>
<tr>
<td>( m_i )</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>( M_i )</td>
<td>61</td>
<td>75</td>
<td>71</td>
<td>123</td>
<td>127</td>
<td>112</td>
<td>163</td>
</tr>
<tr>
<td>( M_i ) (± s.e.)</td>
<td>1,464±540</td>
<td>1,294±269</td>
<td>1,373±508</td>
<td>1,870±490</td>
<td>2,307±443</td>
<td>1,699±266</td>
<td>4,510±1827</td>
</tr>
</tbody>
</table>

We now need to estimate the total number of cattle in Lobamba Sub-region.
Substituting \( n=4 \) an \( N=13 \) in the appropriate formulae (Exploration & description / Formulae) we have:

mean number of animals per homestead = \( \bar{y} = (T_1+T_2+T_3+T_4)/(M_1+M_2+M_3+M_4) = 14.90. \)

estimated total number of cattle in region = \( T = \bar{y} (M_1+M_2+M_3+M_4) \times 13/4 = 24,403 \)

s.e. (\( T \)) = \( \sqrt{13(13-4) [(T_1 - T)^2+...+(T_4 - T)^2] / 4} + 13 (S_1^2+...+S_4^2) / 4) = 4,276. \)

The calculations are shown in CS12Data1

Applying these calculations to all sub-regions we get:

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Population</th>
<th>Sample</th>
<th>Population estimates</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lobamba</td>
<td>13</td>
<td>4</td>
<td>24,403</td>
<td>4,276</td>
</tr>
<tr>
<td>Malandzela</td>
<td>12</td>
<td>7</td>
<td>24,884</td>
<td>4,310</td>
</tr>
<tr>
<td>Mayiwan</td>
<td>20</td>
<td>15</td>
<td>77,510</td>
<td>6,960</td>
</tr>
<tr>
<td>Mbabane</td>
<td>28</td>
<td>15</td>
<td>84,489</td>
<td>9,996</td>
</tr>
<tr>
<td>Ntonjani</td>
<td>13</td>
<td>7</td>
<td>32,897</td>
<td>4,421</td>
</tr>
<tr>
<td>Piggs Peak</td>
<td>14</td>
<td>5</td>
<td>46,526</td>
<td>6,555</td>
</tr>
</tbody>
</table>

Scenario 2

In this scenario we are assuming that we now know the sizes of all dip-tank areas (see below).

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Dip-tanks Population (N)</th>
<th>Sample (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lobamba</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Malandzela</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Mayiwan</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Mbabane</td>
<td>28</td>
<td>20</td>
</tr>
<tr>
<td>Ntonjani</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Piggs Peak</td>
<td>14</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Dip-tank area Population</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lobamba</td>
<td>Buhlungu 105</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Ezulwini 126</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Luphahlane 147</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Mhlambanyatsi 126</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Bhidillili 183</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mdonjane 113</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sigcawini 76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vezl 99</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Makhebelele 61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Luyingweni 87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mambiteni 55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Magwama 106</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Majuba 60</td>
<td></td>
</tr>
</tbody>
</table>

We follow the same calculations as before for estimating \( T_i \)s and \( S_i^2 \)s for each sampled dip-tank area, as shown in 'Cat_no_pop_inf' worksheet of CS12Data2 and starting with Lobamba
Sub-region.

In order to estimate population numbers in the whole sub-region, however, the formulae are different - see Exploration & description / Formulae.

As before, mean number per homestead = \( \overline{Y} = \frac{T_1 + T_2 + T_3 + T_4}{M_1 + M_2 + M_3 + M_4} = 14.90 \)

But now, estimated total number of cattle in region = \( T = \overline{Y} (M_1 + \ldots + M_1) = 20,023 \)

\[ \text{s.e.}(T) = \sqrt{\frac{13(13-4)}{(4-1)} \left[ (T_1 - \overline{Y} M_1)^2 + \ldots + (T_4 - \overline{Y} M_4)^2 \right] + 13(S_1^2 + \ldots + S_4^2) / 4} = 5,327 \]

The results for all sub-regions are as follows:

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Dip-tanks</th>
<th>Population estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Lobamba</td>
<td>13</td>
<td>20,023</td>
</tr>
<tr>
<td>Malandzela</td>
<td>12</td>
<td>22,408</td>
</tr>
<tr>
<td>Mayiwane</td>
<td>20</td>
<td>66,038</td>
</tr>
<tr>
<td>Mbabane</td>
<td>28</td>
<td>73,841</td>
</tr>
<tr>
<td>Ntfonjeni</td>
<td>13</td>
<td>23,462</td>
</tr>
<tr>
<td>Piggs-Peak</td>
<td>14</td>
<td>27,008</td>
</tr>
</tbody>
</table>

Putting the results for the two scenarios together we get:

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Dip-tank areas</th>
<th>Population estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Scenario 1</td>
</tr>
<tr>
<td>Lobamba</td>
<td>13</td>
<td>24,403 ± 4,276</td>
</tr>
<tr>
<td>Malandzela</td>
<td>12</td>
<td>24,884 ± 4,310</td>
</tr>
<tr>
<td>Mayiwane</td>
<td>20</td>
<td>77,510 ± 6,960</td>
</tr>
<tr>
<td>Mbabane</td>
<td>28</td>
<td>84,489 ± 9,996</td>
</tr>
<tr>
<td>Ntfonjeni</td>
<td>13</td>
<td>32,897 ± 4,421</td>
</tr>
<tr>
<td>Piggs-Peak</td>
<td>14</td>
<td>46,526 ± 6,555</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>290,709 ± 15,740</td>
</tr>
</tbody>
</table>

As all sub-regions are sampled there is no further sampling variation to add when calculating the standard error for the grand total. There is, therefore, no additional term to add to the formula for the sampling variance.

The standard error for the total population is thus the square root of the sum of variances for each sub-region. For scenario 1 this is:

\[ \sqrt{(4,276^2 + 4,310^2 + \ldots + 6,555^2)} = 15,740. \]

Similarly the standard error for scenario 2 is 12,756.

We thus have two estimates for the total number of cattle in Hhohho Region. Calculating 95% confidence intervals (2 x s.e.) we obtain:

<table>
<thead>
<tr>
<th>Total</th>
<th>95% confidence range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>290,709 - 259,229 - 322,189</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>232,780 - 207,268 - 258,292</td>
</tr>
</tbody>
</table>
Thus, we have two quite different answers with two confidence intervals that barely overlap.

We shall question later under Findings, implications and lessons learned the assumptions we make with each of the two scenarios.

### Summaries

Applying the formulae from **Scenario 1** to **CS12Data1** for each of the livestock species we obtain the following preliminary results. Firstly, for each sub-region:

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Cattle (Total)</th>
<th>Sheep (Total)</th>
<th>Goats (Total)</th>
<th>Chickens (Total)</th>
<th>Pigs (Total)</th>
<th>Donkeys (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24403</td>
<td>4276</td>
<td>10674</td>
<td>2872</td>
<td>28777</td>
<td>4435</td>
</tr>
<tr>
<td>2</td>
<td>24884</td>
<td>4310</td>
<td>5182</td>
<td>5227</td>
<td>26267</td>
<td>5823</td>
</tr>
<tr>
<td>3</td>
<td>77510</td>
<td>6960</td>
<td>48701</td>
<td>4747</td>
<td>65009</td>
<td>3255</td>
</tr>
<tr>
<td>4</td>
<td>84489</td>
<td>9996</td>
<td>75052</td>
<td>11175</td>
<td>161996</td>
<td>64629</td>
</tr>
<tr>
<td>5</td>
<td>32897</td>
<td>4421</td>
<td>3479</td>
<td>1510</td>
<td>13391</td>
<td>1847</td>
</tr>
<tr>
<td>6</td>
<td>46526</td>
<td>6555</td>
<td>50866</td>
<td>8376</td>
<td>71229</td>
<td>3392</td>
</tr>
</tbody>
</table>

and, secondly, for the region as a whole:

<table>
<thead>
<tr>
<th>Species</th>
<th>Estimated population total</th>
<th>Standard error</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cattle</td>
<td>233,000</td>
<td>12,800</td>
<td>208,000 to 258,000</td>
</tr>
<tr>
<td>Sheep</td>
<td>24,300</td>
<td>4,290</td>
<td>15,900 to 33,000</td>
</tr>
<tr>
<td>Goats</td>
<td>197,000</td>
<td>13,400</td>
<td>171,000 to 223,000</td>
</tr>
<tr>
<td>Chickens</td>
<td>319,000</td>
<td>68,400</td>
<td>184,000 to 453,000</td>
</tr>
<tr>
<td>Pigs</td>
<td>20,600</td>
<td>2,920</td>
<td>14,900 to 26,000</td>
</tr>
<tr>
<td>Donkeys</td>
<td>3,960</td>
<td>1,090</td>
<td>1,820 to 6,090</td>
</tr>
</tbody>
</table>

Note that two homesteads with large numbers of sheep recorded for them have been removed for the calculation of population estimates for sheep (see **CS12Doc**).

Applying the formulae from **Scenario 1** to **CS12Data1** for each of the livestock species we obtain the following preliminary results. Firstly, for each sub-region:
and, secondly, for the region as a whole:

<table>
<thead>
<tr>
<th>Species</th>
<th>Estimated population total</th>
<th>Standard error</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cattle</td>
<td>233,000</td>
<td>12,800</td>
<td>208,000 to 258,000</td>
</tr>
<tr>
<td>Sheep</td>
<td>197,000</td>
<td>4,290</td>
<td>159,000 to 233,000</td>
</tr>
<tr>
<td>Goats</td>
<td>319,000</td>
<td>36,400</td>
<td>255,000 to 380,000</td>
</tr>
<tr>
<td>Chickens</td>
<td>20,600</td>
<td>2,920</td>
<td>14,900 to 26,000</td>
</tr>
<tr>
<td>Pigs</td>
<td>3,960</td>
<td>1,090</td>
<td>1,820 to 6,090</td>
</tr>
<tr>
<td>Donkeys</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note that two homesteads with large numbers of sheep recorded for them have been removed for the calculation of population estimates for sheep (see CS12Doc).

Please note that these are preliminary estimates. At this early stage, however, we can say that chickens occur, as one might expect, in the greatest numbers, followed by cattle and goats. Other species exist in much smaller numbers.

We have, however, made various assumptions.

For example, we have assumed for the purpose of this case study that dip-tank areas were selected completely at random and that there was no stratification. Correct population estimates will need to be obtained taking into consideration the precise structure of the sampling design.

There are also a couple of somewhat large standard errors appearing for chickens (shaded yellow on previous page). The source of this problem needs to be investigated by studying extreme values in the raw data as we did for sheep.

The standard errors for cattle and goats are 5-7% of the mean. This is reasonable. However, for sheep, pigs and donkeys, they are much higher. This is understandable because of the large numbers of homesteads without these species.

Finally, note how the results in the second table have been presented. Numbers are reported with 3 significant figures. These make the results very much easier to follow than those presented in the first table. This is an important point for reporting.

**Findings, implications and lessons learned**

There are many things of a statistical nature to learn from the foregoing that we shall note here.
It is most important to record precisely how sampling was done. Any choice of sampling that is strictly non-random affects the ability to estimate population numbers without bias.

- We have obtained two population estimates, firstly, knowing the numbers of homesteads in all dip-tank areas and, secondly, knowing only those in dip-tank areas that were sampled. It follows that if we don’t know how the dip-tank areas were selected we cannot say which method, if either, is correct.

- Sampling dip-tanks in such a survey is helped by taking into account the different agro-ecological zones, the proximity to urban areas and the livestock densities in order to form more homogeneous strata within which dip-tank areas can be sampled.

- Stratification has the further benefit of reducing standard errors as it puts similar types of dip-tank areas together, and hence reduces the among dip-tank area component of the variance.

- Although stratification was used in this survey, we have deliberately ignored it so as not to overcomplicate the calculations.

- It was observed that the distribution of cattle per homestead was skewed. It is possible that a logarithmic transformation might be applied to the raw data in order to improve the precision of the population estimate. One of the study questions addresses this.

- Two different formulae have been used for calculating standard errors of population totals:

\[
S_T = \sqrt{\frac{\left( (T_1 - \bar{T})^2 + \cdots + (T_n - \bar{T})^2 \right) / (n-1)}{}}
\]

\[
S_T = \sqrt{\frac{\left( (T_1 - \bar{Y}_{M_1})^2 + \cdots + (T_n - \bar{Y}_{M_n})^2 \right) / (n-1)}{}}.
\]

We have applied the first formula to Scenario 1 (equal probability) and the second to Scenario 2 (probability proportional to dip-tank area size).

We can also apply the second formula to Scenario 1, (equal probability of sampling) where \(M_i\) can be assumed to be the value of any ancillary variable associated with a dip-tank area.

This could be from census information such as total number of homesteads (including those without livestock), dip-tank area size in hectares or human population, each of which might correlate with \(T_i\).

Thus, if the number of homesteads with livestock in the sampled areas is unknown then one should look for other national census statistics that might be available.
As we have seen earlier for the two scenarios use of the second formula with ancillary data $M_i$ in general reduced the standard error. This is again shown in the table below. All except the standard error for Lobamba were reduced when the second formula with $M_i = \text{number of homesteads with livestock}$ was used.

Use of ancillary data is useful if $M_i$ is correlated with $T_i$.

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Dip-tank areas</th>
<th>Population</th>
<th>Sample</th>
<th>Standard errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Formula 1</td>
</tr>
<tr>
<td>Lobamba</td>
<td>13</td>
<td>13</td>
<td>4</td>
<td>4,276 (0.18)$^a$</td>
</tr>
<tr>
<td>Malandzele</td>
<td>12</td>
<td>12</td>
<td>7</td>
<td>4,310 (0.17)</td>
</tr>
<tr>
<td>Mayiwane</td>
<td>20</td>
<td>20</td>
<td>15</td>
<td>6,960 (0.09)</td>
</tr>
<tr>
<td>Mbabane</td>
<td>28</td>
<td>23</td>
<td>15</td>
<td>9,996 (0.12)</td>
</tr>
<tr>
<td>Ntfonjeni</td>
<td>13</td>
<td>13</td>
<td>7</td>
<td>4,421 (0.13)</td>
</tr>
<tr>
<td>Pigg-Peak</td>
<td>14</td>
<td>14</td>
<td>5</td>
<td>6,555 (0.14)</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>94</td>
<td>58</td>
<td>15,740</td>
</tr>
</tbody>
</table>

$^a$ ratio of standard error to population mean

Another point to note in the above table is how the standard error as a proportion of the population mean decreases as the sampling ratio of dip-tank areas increases. Contrast the ratio of 0.09 in the table above when 75% of dip-tank areas were sampled for Mayiwane Sub-region compared with the ratio of 0.18 when 31% of dip-tank areas were sampled in Lobamba Sub-region.

The variation among dip-tank areas is greater than the variation among homesteads within dip-tank areas. This is generally true in cluster sampling - variation increases as one moves up the hierarchy.

Thus, one can see in CS12Data2 (worksheet Cat\_pop\_inf) that the contribution of the variance from dip-tank areas to the standard error of the population mean (e.g. 22,928,938 for Lobamba) is greater than that for homesteads (5,449,984 for Lobamba).

For Mayiwane, however, the larger proportion of dip-tank areas sampled has resulted in a smaller among dip-tank area variance (17,539,279) despite the larger variance that occurred within dip-tank areas (27,084,384).

It can be seen how important it is, therefore, to ensure that there is adequate replication at the upper layers of a clustered hierarchy when planning a sampling frame.
The most efficient form of sampling in terms of precision of population estimates is to sample proportional to size. Thus, the 60 dip-tank areas to be sampled would ideally be distributed so that similar proportions in each sub-region are sampled relative to the total number of dip-tank areas in the sub-regions.

The 600 or so homesteads within the Hhohho Sub-region would then be distributed across selected dip-tank areas proportional to the numbers of homesteads within the areas. This would be preferable to that of choosing a constant number, as was the intention in this case study, for each dip-tank area.

Finally, it should be noted that different statistical software that can be used to carry out calculations of standard errors of population estimates derived from cluster sampling sometimes result in slight differences in results, and may even differ from those derived here.

More than likely this may be due to slight differences in the construction of the formulae. One should nevertheless verify that this is indeed the reason and not a result of misunderstanding of how the calculations have been done.

Study questions

1. Write a summary of the results given for the six livestock species suitable for a scientific audience. Assume in your summary that no particular stratification was applied and that dip-tank areas were not selected at random but chosen to give a widespread representation of a sub-region.

2. Repeat the calculations in CS12Data1 for Scenario 1 for cattle but replace the formula \( S_T^2 = [(T_1 - \bar{T})^2 + \ldots + (T_n - \bar{T})^2] / (n-1) \) by \( S_T^2 = [(T_1 - Y \cdot \bar{M}_1)^2 + \ldots + (T_n - Y \cdot \bar{M}_n)^2] / (n-1) \). Compare the confidence limits with those obtained for Scenario 1 in this case study.

3. The numbers of the different livestock species per homestead are not normally distributed. Apply the logarithmic transformation to the numbers of animals per homestead and repeat the calculations in CS12Data1 and CS12Data2 for cattle. Express the population results in terms of an estimated total together with confidence
limits. Comment on the results.

4. Assume that strata are selected in sub-region $i$ that define areas of different livestock density. Assume that $m_{ij}$ homesteads are now sampled from each stratum and that each stratum contains $M_{ij}$ homesteads ($j=1...r$). Write down the algebraic formulae that you will now need to estimate the total number of cattle in sub-region $i$.

5. The researcher reports that the Buhlungu dip-tank area in Lobamba Sub-region was selected purposively because of a particular breed of cattle known only to occur there. The other dip-tank areas were chosen at random. Decide how you would use this information to revise the calculation of the population total and standard error for Lobamba. Obtain new population estimates.

6. In view of the few donkeys kept by farmers in Swaziland it is decided to analyse the proportion of homesteads owning cattle per dip-tank area rather than the number of donkeys owned per homestead. This then reduces to a single-stage analysis at the sub-region level. Write down the formulae you will need to calculate the population average and its standard error. Apply this formula to Malandzela Sub-region. Multiply the population average and its standard error by the number of homesteads to get an estimate of the sub-region confidence range. Compare with that already obtained.

7. It is decided to select instead 6 dip-tank areas within the Lobama sub-region at random so that the chance of each dip-tank area being selected is proportional to the numbers of homesteads in that dip-tank area. Describe how you would do this.

8. Check the raw data for chickens and delete any homesteads with extremely high numbers of chickens that might be considered to be outliers. Repeat the calculations of the two population estimates for chickens. Do they seem more reasonable?

9. You reach a dip-tank area with your assistant and are ready to select 10 homesteads at random. Describe how you might select the sample.

10. Stratification of dip-tank areas can be based on such factors as livestock density, ecological zone and proximity to a commercial farm. It is required to select 6 dip-tank areas from the following dip-tank areas for Lobamba Sub-region. Decide how you would stratify the dip-tank areas and which factors you would use to define the strata. Select dip-tank areas from within these strata at random to make up the six required.

<table>
<thead>
<tr>
<th>Dip-tank area</th>
<th>Livestock density</th>
<th>Ecological zone</th>
<th>Close to commercial farm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buhlungu</td>
<td>Medium</td>
<td>A</td>
<td>Yes</td>
</tr>
<tr>
<td>Ezulwini</td>
<td>Low</td>
<td>A</td>
<td>No</td>
</tr>
<tr>
<td>Luphahlane</td>
<td>Low</td>
<td>A</td>
<td>No</td>
</tr>
<tr>
<td>Mhlambanyatsi</td>
<td>Medium</td>
<td>B</td>
<td>No</td>
</tr>
<tr>
<td>Bhidillili</td>
<td>High</td>
<td>C</td>
<td>No</td>
</tr>
<tr>
<td>Mdonjane</td>
<td>High</td>
<td>B</td>
<td>Yes</td>
</tr>
<tr>
<td>Sigcawini</td>
<td>Medium</td>
<td>B</td>
<td>Yes</td>
</tr>
<tr>
<td>Vezi</td>
<td>High</td>
<td>A</td>
<td>No</td>
</tr>
<tr>
<td>Makhebele</td>
<td>High</td>
<td>A</td>
<td>No</td>
</tr>
<tr>
<td>Luyengweni</td>
<td>Medium</td>
<td>A</td>
<td>No</td>
</tr>
<tr>
<td>Mambiteni</td>
<td>Low</td>
<td>A</td>
<td>No</td>
</tr>
<tr>
<td>Magwama</td>
<td>Medium</td>
<td>B</td>
<td>No</td>
</tr>
<tr>
<td>Majuba</td>
<td>High</td>
<td>B</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Acknowledgements

The survey in Swaziland was conducted by the Ministry of Agriculture and Co-operatives (Department of Veterinary and Livestock Services), in collaboration with the University of Swaziland. We give sincere thanks to all those who were involved from the planning stage to the data collection and data capture and cleaning. Also, many thanks to the farmers in the Hhohho region for their kindness and time they took to answer the questions in the survey questionnaire.

We are grateful to the Department of Veterinary and Livestock Services for permission to use these data from Hhohho Region for this case study. However, readers should appreciate that we have ignored the sampling frame used to select dip-tank areas. Thus, the population estimates produced will be incorrect and hence must not be quoted. Users are welcome to use the data but, for the same reason, may not publish any findings.

The first author made his contribution to the work presented in this case study while he was a visiting Research Fellow at the School of Applied Statistics, University of Reading. He would like to sincerely thank the School and the University for hosting his visit and for all the support and assistance afforded to him during the visit. In particular, he would like to give sincere thanks to the Head of School, Dr. Derek Pike, for supervising the work.
The first author visited the University of Reading through a Commonwealth Fellowship Award offered by the Commonwealth Scholarship Commission in the United Kingdom under the Commonwealth Scholarship and Fellowship Plan. He would like to sincerely thank the Commission for the Award, and to the British Council, which was responsible for all welfare aspects of the Award.

During the tenure of the first author's Commonwealth Fellowship Award in Britain, he was on paid Sabbatical Leave with his employer, the University of Swaziland. He would like to sincerely thank the University of Swaziland for approving his Sabbatical Leave and for the financial support the University afforded him to have his family join him in the United Kingdom during his Fellowship at the University of Reading.

Source: Thalile Seipone Sowane