



Welsh Agricultural Statistics – agricultural statistics using StatsWales: an introduction

1 August 2017
SB 41/2017

Purpose of bulletin

The bulletin announces the release of some trial agricultural data on Welsh Government's flexible access tool [StatsWales](#).

The data are estimates from the June 2014 Welsh Agricultural Survey. The results show estimates for the main land and livestock variables for subgroups of farms across Wales. There are a number of different subdivisions based on region, farm type, farm economic size, numbers of various animals on a farm and so on.

The main aim is to test to what extent this sort of release of data is valuable to users. We seek feedback on the details of which breakdowns are shown and the categories within them. However, the range of results that can be released in this way needs to be a compromise between the level of detail desired and maintaining a reasonable level of data quality.

If the approach is agreed to be useful then, with suitable modifications, it can become part of the annual release of data from the Welsh Agricultural Survey. The intention would be to be a full series of data including previous surveys and adding new surveys as they become available.

About this bulletin

Announcing a new set of StatsWales cubes with data about agriculture.

The purpose of this bulletin is to check with users that this way of presenting data is useful and to seek comments on the detail of what is included.

In this bulletin

Introduction to StatsWales	2
Overview of the data released	3
Variables released	4
StatsWales cubes	5
Data quality	7
Disclosure	8
Example analysis	8
Further information	11
Contact details	13

Introduction to StatsWales

StatsWales is a tool for flexible access to data giving users more choices than would be possible in simple tables, spreadsheets or other traditional formats. This is the first time that it has been used for agricultural datasets.

The flexibility of StatsWales comes at the cost of needing a little training in how the application works. This is not a large overhead as the application is quite intuitive once a few key concepts have been understood. There is a full introduction to [StatsWales on-line](#) including a tutorial to show how it can be used.

The building block for StatsWales is the “data cube”. For the agricultural data the cubes all have three dimensions:

- A common set of land and livestock categories.
- The survey year. Note that in this set of trial data only a single year is given. If users find the approach useful new data will be added each year and the historic data will also be added over time.
- A variable that splits the Wales level totals into a number of categories. For individual cubes this may be regions, farm economic size, area of land on a farm, number of animals on a farm and so on.

The package allows users to manipulate the cubes. We can examine the distribution of any variable in the common set by, for example, the number of sheep on a farm. In this example the first thing is to examine the number of farms in each size band together with the number of sheep. This analysis has previously been included in the compendium publication “Welsh Agricultural Statistics”. The new release allows you to also look at the distribution of, say, land or cattle by the number of sheep on a farm giving more depth to the analysis and starting to look at the relationships between different agricultural activities.

Thus the benefit is that users can specify tables themselves. The results can be exported to spreadsheets. If there are queries that users will require regularly then these can be saved so that they do not have to be set up each time.

Overview of the agricultural survey data released

A number of StatsWales agricultural survey data cubes are released. Key points are as follows.

- Estimates are from the June 2014 Welsh Agricultural Survey
- There is a common list of land, livestock and economic variables that is reported on all the StatsWales data cubes. This is not a complete list of all the variables captured on the June Agricultural survey but is tailored to be a good compromise between the level of detail in the variables with having sufficient farms in each cell of the cube.
- Each individual cube breaks down the Wales level data into a number of categories. These may be regions of Wales, the dominant activity on a farm, the economic size of the land, or bands based on the numbers of animals or areas of land on the farm.

These data cubes extend the results that have already been released.

There is a [spreadsheet with the Wales level totals for all variables collected in the survey from 1998 to 2016](#). The Wales totals in the StatsWales cubes agree with the previously published totals.

The [Welsh Agricultural Statistics](#) publication has a number of tables by farm type, numbers of livestock and areas of land. The [StatsWales cubes](#) extend these tables by including all the agricultural variables in the common list.

For example, Welsh Agricultural Statistics has a table showing the distribution of farms and dairy cows by the number of dairy cows on a farm. The StatsWales cube by number of dairy cows allows the user to see distribution of other cattle (beef cows, calves and others) by the number of dairy cows on a farm. Thus we can get a picture of the mix of cattle on farms with different size dairy herds. Similarly the cube allows an initial view of the mix of land and other livestock types. Note that the categories for the number of dairy cows on a farm are different for the StatsWales and the Welsh Agricultural Statistics tables.

It is not feasible to produce all possible cubes. The intention is to produce a set of key cubes that can be used as a starting point for a wide range of users. The set being released now is a first attempt at this key set. It has been designed as a compromise between having a good range of detail while also maintaining a reasonable level of data quality.

Variables released

There is a common set of key variables that is reported for all the StatsWales cubes. This is not the entire set of variable collected through the Survey. Comments are invited on the usefulness of the common list – within the restriction of needing to ensure adequate numbers of farms in all categories.

Count of farms

Simple count of the number of farms. An important feature of released data is that the farm counts can be linked with the amount of land and livestock on the group of farms. The intention is that this will help stop the practice of only reporting farm counts (or share of farms) without considering the wider implications.

Number of responses

The number of farms for which there was an actual response to the survey.

Land on farms

- Permanent pasture not resown within the last 5 years
- New grassland resown within the last 5 years. This may be recently resown permanent pasture but can also be grass used as part of an arable rotation.
- Sole rights rough grazing. Excludes the use of common land. Note that the choice of reporting either permanent pasture or rough grazing is at the discretion of the farmer.
- Cereals – mainly wheat and barley but includes other cereals
- Stockfeed – crops for feeding to livestock includes maize and other crops specifically for livestock.
- Other crops – any other arable or horticultural crops not in cereals or stockfeed. Largest single crop is oil seed rape. Includes potatoes and horticultural land use.
- Woods and others – includes farm woodland, also ponds, paths, buildings and any non-agricultural land on the farm (for example a permanent caravan park or a solar farm)

Livestock

- Dairy cows – calved females of a dairy breed
- Beef cows – calved females of a beef breed
- Calves under 1 year – all cattle aged under 1 year
- Other cattle – female cattle who have not had a calf aged at least 1 year and all males aged at least 1 year
- Sheep – all ewes, lambs, rams and others
- Poultry – chickens for meat, eggs or breeding and other poultry such as turkeys, ducks or geese.
- Pigs – all pigs for fattening or breeding

Economic size measures

These are proxy measures that are based on weighted sums of the crop areas and livestock numbers. Each is a common method for the EU with weights calculated for the UK. The proxy measures are based on what would be expected on an average farm with a certain number of livestock and area of land. At the farm level the unknown actual value will vary around this average.

The following proxies are reported:

- Standard Output (SO) – a proxy for turnover on the farm
- Standard Gross Margin (SGM) – a proxy for profit on the farm
- Standard Labour Requirement (SLR) – a proxy for the amount of work that needs to be done on a farm. This is converted into the number of “full time equivalent” workers needed on a farm. What it will not show is how this requirement is met through full time, part time, casual or contract working

The measures are usually used at the aggregate level to compare the contribution from different activities, farm types, size, regions and so on. They are also useful to classify farm economic size in various ways. In general the focus for analysis should be the relative values in preference to the absolute values.

Farm labour

The release does not have estimates for the number of farmers and farm workers. This is due to the continuing issues with the quality of the farm labour data from the survey. Total farm labour estimates for Wales as a whole are available.

For those interested in farm labour, the variables include the Standard Labour Requirement described above.

Individual StatsWales cubes

Each cube will show the full common variable list for each year available. The individual cubes divide the set of farms in Wales into a small set of categories. In the initial release of data the following categories are used.

- Farm type – showing the dominant (but not necessarily only) activity on a farm. Two sets of farm types are used. One is a summary of the usual Robust Farm Types commonly used in UK agricultural statistics. The second is a version of farm types that is focused specifically on Wales.
- Farm economic size – showing economic size distributions for the three proxy measures Standard Gross Margin, Standard Output and Standard Labour Requirement.
- Local authority area. The full set of local authorities are included to demonstrate why showing estimates for all 22 authorities is not realistic. The question is which authorities should be grouped. In the StatsWales table the authorities are grouped into the current seven agricultural regions.
- By area of land on a farm

- By number of cattle on a farm
- By number of dairy cows on a farm
- By number of beef cows on a farm
- By number of sheep on a farm
- By total area of crops on a farm

The set of cubes released gives a wide coverage of themes covered by the Agricultural Survey. It is not exhaustive and cannot be optimal for all users or for all uses. They are intended to act as a general reference and a starting point for further more detailed analyses. With that in mind we would like users to comment on how well they think this meets their needs and any modification that would make them more useful. There are three key things that we would like users to comment on

- Is the range of cubes about right? Is there anything missing or anything that could be omitted?
- Are the various bands used to show the distribution about right? We will never be able to have a set of bands that meet all requirements for all users. The aim is to provide a good starting point that is helpful for as many uses as possible.
- The presentation of regional values is a particular issue. Some level of aggregation from local authorities is clearly necessary. The question is what level of aggregation gives the best balance between the detail and the quality of the data?

Data quality

An important limitation on what should be released is the need that the estimates should remain reasonably reliable. There comes a point with any survey estimates where any genuine information is swamped by the noise from the various sources of error in the survey.

Key points

- The results are estimates with a degree of uncertainty around them.
- The reliability of the estimates depends mainly on the number of farms on which the estimates are based.
- The registration of farms is not perfect. While a commercial farm has a strong incentive to register there is little incentive to deregister when a farm stops activity. This can lead to farms that have closed remaining in the survey population, inflating the survey estimates.
- The survey only goes to a sample of the registered farms and not all those selected respond. Estimates are required for farms that are either not selected or that do not respond. Sample selection is weighted so that economically larger farms are sampled more regularly than smaller farms. This is to give good coverage of the land and livestock in Wales while limiting the sample size.
- For users there must always be a trade off between the level of detail wanted and the data quality of the estimates. The assumption that increasing the level of detail is always beneficial ignores this trade off which is an important element of judging whether the data may be fit for the purposes to which users wish to put it.

Note that the data for cattle is not collected from the survey. These figures are taken from the registrations with the Cattle Tracing System. It is still important to know if a farm with cattle was a response or not if we want to relate the number of cattle on a farm to other activities on the farm. For example, to calculate the cattle per hectare or to find farms that have both cattle and sheep.

In the released data there are flags set to warn the user when estimates are based on small numbers of farms. These show as note markers on the StatsWales tables. Two levels of warning area given:

- Cells based on less than 25 responses are flagged as unreliable
- Cells with between 25 and 49 responses are flagged as a warning

So that the results add up correctly and can be used in calculations the results are not suppressed. However, the users can see that there are data quality issues and treat values with a degree of caution.

Disclosure

Where there are very small numbers of farms in a particular cell of a cube there can also be a risk of revealing information about individual farms – statistical disclosure. For example, if there was only a single large poultry unit in a particular region you might be able to make a good estimate of how many poultry it had from the data.

For this trial release the only treatment for potential disclosure has been in the selection of the categories. The categories within each cube are selected so that there are reasonable numbers of farms in each.

As mentioned in the [data quality](#) section, the individual cells are flagged when they are based on particularly small numbers of farms. The threshold for a warning flag is relatively high, being based on how many farms are needed from a data quality point of view. This will be much higher than the number below which disclosure becomes an issue.

This approach is based on three main considerations

- The nature of the survey, and farm registrations, mean that actual identification of individual farms is unlikely. The issues with non-response, imputation and the allocation to categories reduce the likelihood of an individual farm actually being identified.
- The nature of farms means that their presence in an area is usually obvious to those in the area. How much extra information is revealed using the land and livestock estimates from the survey is questionable, even from actual responses rather than estimated results.
- Suppression of some results within some cubes would limit how much use can be made of the results.

Users views are invited on whether this is an acceptable approach or if more robust disclosure control is required.

Example analysis

An example of how the data might be used is probably helpful.

Using the StatsWales cubes we can examine the distribution of key agricultural variable by the economic size of the farm. This could use the cube based on the Standard Output on a farm (Standard Output being a proxy measure for turnover). Extracting the data from StatsWales we can produce a simple table showing the share of key activities (land areas and livestock numbers) by the Standard Output size bands. Table 1 shows the results for this for the 2014 data. This gives a simple way to see the spread of land and livestock across the size bands. Potentially if the cubes are expanded to show a range of years it would be quite simple to look at, for example, the way that the share of sheep on farms with Standard Output of at least €200,000 has change over time.

To produce the table in this example, the relevant view is selected in StatsWales and the data downloaded, using the “spreadsheet” option, and the final presentation of these results was done using MS-Excel. Note it is also possible to create a permanent machine readable link to the data as explained in the [StatsWales help](#) section.

Table 1**Share of key agricultural variables by Standard Output on a farm, 2014**

	Standard Output on a farm (€'000)					Total (=100%)
	Under 25	25 - 50	50 - 100	100 - 200	At least 200	
Farms	58.8%	12.7%	11.9%	8.9%	7.8%	24,722
Standard Output	6.2%	6.9%	12.8%	18.6%	55.5%	1,686,793
Land on farms						
Perm pasture	17.2%	14.8%	21.4%	22.9%	23.7%	1,047,916
Rough grazing	9.8%	13.8%	26.9%	26.2%	23.2%	257,055
New grass	8.0%	9.8%	17.3%	22.8%	42.0%	152,571
Crops & hort	3.6%	6.1%	14.1%	22.2%	53.9%	86,728
Woods & others	26.6%	14.1%	19.6%	21.1%	18.6%	87,094
Total	15.0%	13.6%	21.4%	23.3%	26.7%	1,631,364
Cattle						
Dairy cows	0.3%	0.4%	1.5%	10.5%	87.4%	234,305
Beef cows	9.3%	15.4%	26.5%	29.5%	19.3%	167,799
Other cattle	5.5%	9.3%	16.7%	21.6%	47.0%	700,664
Total	5.0%	8.3%	15.0%	20.4%	51.3%	1,102,768
Other livestock						
Sheep	6.9%	11.5%	24.5%	32.1%	24.9%	9,738,871
Poultry	1.1%	0.2%	0.5%	1.4%	96.8%	8,997,200
Pigs	19.8%	5.2%	7.2%	6.8%	60.9%	28,363

Source: estimates from Welsh Agricultural Survey, June 2014

In Table 1 land area totals are in hectares and livestock totals are headcounts. The total Standard Output is in thousands of euros.

Further information

Welsh Government produces a range of regular agricultural statistics releases and also one off reports from time to time. These can be found on the “[farming](#)” sub topic on the “[statistics and research](#)” section of the Welsh Government website.

The releases include

- [Annual release of land and livestock](#) data at the Wales level. This release has a spreadsheet of estimates for all land and livestock variables from 1998 to 2016.
- [Welsh Agricultural Statistics](#). Compendium volume.
- [Agricultural Small Areas](#) bulletin. Description of the data quality issues. Also a spreadsheet with estimates for most variables at the level of the 235 Agricultural Small Areas. When a variable in an individual Small Areas is potentially disclosive then the value has to be suppressed in this spreadsheet.
- [Farm incomes](#) and [Agricultural Aggregate Account](#). Releases with information about farming finance.

StatsWales links

- [Statwales homepage](#)
- [Statwales help](#)
- [Agriculture on StatsWales](#)

National Statistics status

The [United Kingdom Statistics Authority](#) has designated these statistics as National Statistics, in accordance with the Statistics and Registration Service Act 2007 and signifying compliance with the [Code of Practice for Official Statistics](#).

National Statistics status means that official statistics meet the highest standards of trustworthiness, quality and public value. All official statistics should comply with all aspects of the Code of Practice for Official Statistics. They are awarded National Statistics status following an assessment by the UK Statistics Authority's regulatory arm. The Authority considers whether the statistics meet the highest standards of Code compliance, including the value they add to public decisions and debate.

It is Welsh Government's responsibility to maintain compliance with the standards National Statistics. If we become concerned about whether these statistics are still meeting the appropriate standards, we will discuss any concerns with the Authority promptly. National Statistics status can be removed at any point when the highest standards are not maintained, and reinstated when standards are restored.

Well-being of Future Generations Act (WFG)

The Well-being of Future Generations Act 2015 is about improving the social, economic, environmental and cultural well-being of Wales. The Act puts in place seven well-being goals for Wales. These are for a more equal, prosperous, resilient, healthier and globally responsible Wales, with cohesive communities and a vibrant culture and thriving Welsh language. Under section (10)(1) of the Act, the Welsh Ministers must (a) publish indicators ("national indicators") that must be applied for the purpose of measuring progress towards the achievement of the Well-being goals, and (b) lay a copy of the national indicators before the National Assembly. The 46 national indicators were laid in March 2016.

Information on indicators and associated technical information - [How do you measure a nation's progress? - National Indicators](#)

Further information on the [Well-being of Future Generations \(Wales\) Act 2015](#).

Statistics in this release could also provide supporting narrative to the national indicators and be used by public services boards in relation to their local well-being assessments and local well-being plans.

Further details

The document is available at:

<http://gov.wales/statistics-and-research/welsh-agricultural-statistics/?lang=en>

Next update

Not a regular output

We want your feedback

We would appreciate users comments, in Welsh or English, on the issues in this bulletin or general agricultural statistics matters.

e-mail: stats.agric@gov.wales

Telephone: 0300 025 6822

Post:

Agriculture and rural affairs statistics

Knowledge and Analytical Services

Welsh Government

Cathays Park

Cardiff

CF10 3NQ

Open Government Licence

All content is available under the [Open Government Licence v3.0](#), except where otherwise stated.

