

## An overview of the cereal usage data collection and quality assurance processes

### The cereal usage surveys

Following a period of public consultation, a partnership between the Defra Statistics Team and AHDB was formed in 2017. Since the start of 2018, AHDB have taken on the responsibility for collecting, validating and publishing the following cereal usage datasets:

Dataset	Collection frequency	Published
GB animal feed production	Monthly	Yes
GB integrated poultry units feed production	Monthly	Yes
UK cereals usage by flour millers and bioethanol producers	Monthly	Yes
UK cereals usage by brewers, maltsters and distillers	Monthly	Yes
UK cereals usage by oat millers	Quarterly	Yes
UK cereals usage in cereal breakfast foods	Quarterly	No*
UK cereal usage by rye millers	Annual	No*

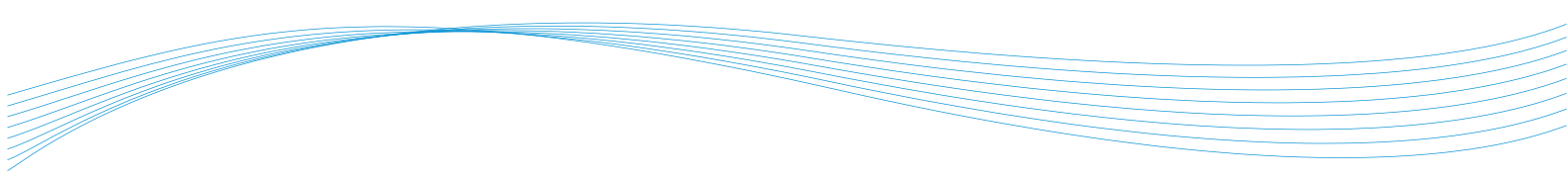
\*Collected for the purposes of the UK cereals supply and demand estimates only

Cereal usage data is published under on two pages on the AHDB website: [UK Human and Industrial \(H&I\) cereal usage](#) and [GB animal feed production](#). This usage data is used to inform the [UK cereals supply and demand estimates](#) which are also delivered on behalf of Defra.

### Data collection and validation process

AHDB survey companies who process raw cereals from different sectors on a monthly, quarterly and annual basis (survey dependant). These sectors include animal feed producers, integrated poultry units, flour millers, oat millers, starch manufacturers, bioethanol producers, rye millers, cereal breakfast food manufacturers and brewers, maltsters and distillers. Within each cereal usage survey, multiple datapoints are collected. These datapoints are aggregated to form data categories (e.g. wheat usage, pig feed production) and, if appropriate, these categories are published on the AHDB website, as part of the datasets listed in the table above.

The below flow diagram provides an overview of the data collection and validation process. All data is handled by a team of trained analysts at AHDB who, using their expertise, validate it through several checks.



**Key:**

**Blue:** Data collection processes

**Orange:** Data quality checks by aggregated category

**Grey:** Data quality checks by company return

**Green:** Creation and distribution of final datasets



## **Further detail on the quality assurance processes carried out by company and data category**

To deliver high-quality data for levy payers and industry, AHDB follow standardised operating procedures and take several vigorous steps to validate and aggregate data collected as part of the cereal usage surveys. The team follow a multi-tiered data validation process which involves both manual and system checks on individual company returns and aggregated data categories. For more information on companies surveyed and categories published, please read above.

Data is quality assured at company level with the following checks carried out:

- Data Analysts check each individual company data returns using custom built excel models to look at year-on-year comparisons and review usage and production trends. Where appropriate, balancing checks take place on stock numbers and usage.
- Data Analysts query any large changes, trend changes and stock adjustments. This will result in data either being accepted or resubmitted.
- Where errors are identified, data providers are asked to resubmit the survey return and the Data Analysts will liaise with them to ensure any back data impacted is also rectified.

Data is quality assured at data category level with the following checks carried out:

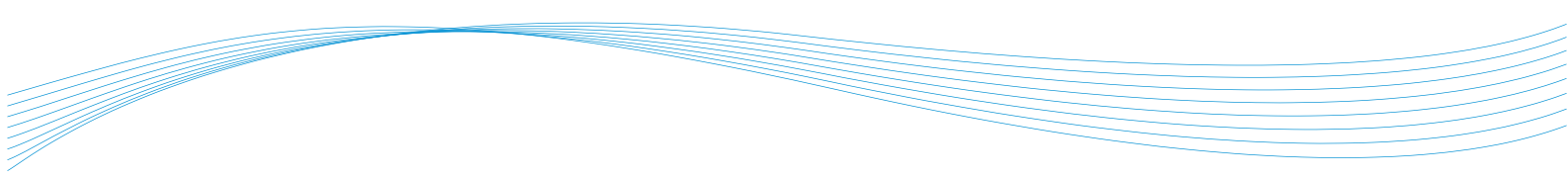
- Data Analysts run reports to create the data categories and check each data category passes the confidentiality thresholds for publication, meaning it is appropriately aggregated and anonymised. These thresholds are based on the number of contributors and their percentage representation.
- Data Analysts review each data category to assess year-on-year and month-on-month changes.
- Data Analysts investigate any large changes or changes in trend using custom built excel models. Company movements are compared side by side, to understand trends within the data categories and to identify key drivers of change. Data Analysts review the queries undertaken as part of the company data quality checks and, where necessary, make further queries with data providers.
- Data Analysts write comprehensive notes for each data category which are passed to a Lead Data Analyst or Head of Department for sign off. This involves a full review of each data category including whether it has met the confidentiality thresholds for publication. Any further queries are flagged to the Data Analyst to resolve before the final datasets are created for publication.

## **Notification of data revisions**

On occasion, revisions need to be made to published cereal usage data, for example when updated data is received from contributors. In this instance, a notification is made on the 'Disclaimer and notes' tab on the downloadable dataset. All notes are dated and regularly reviewed to ensure the most recent changes are clear.

## **Notification of upcoming publication dates**

A cereal usage calendar is published on the AHDB website. This calendar details the number of weeks in a period and a publish date.



### **Continually improving data quality**

AHDB are committed to producing datasets which have the best possible coverage and are representative of the sector. The team continually monitor structural changes in the industry. Where new companies are identified, the team prioritise contacting them for inclusion in the sample and ensure their data is included within the published datasets at the earliest possible opportunity. Where appropriate, back data is also requested to allow accurate historical comparisons.

