Baby names QMI

Quality and methodology information for baby names statistics produced for England and Wales.

Contact:
Kanak Ghosh
Health.Data@ons.gov.uk
+44 1329 444110

Release date:
5 October 2022

Next release:
To be announced

Table of contents

1 . Output information

- National Statistic: yes

- Frequency: annual

- How compiled: based on third party data

- Geographic coverage: England and Wales

- Last revised: October 2022

2 . About this Quality and Methodology Information Report

This quality and methodology report contains information on the quality characteristics of the data – including the European Statistical System five dimensions of quality (PDF, 2.99MB) – and the methods used to create it.

The information in this report will help you to:

- understand the strengths and limitations of the data

- learn about existing uses and users of the data

- help you to decide suitable uses for the data

- reduce the risk of misusing data

3 . Important points about baby names statistics

- Baby name statistics are compiled from first names recorded when live births are registered in England and Wales as part of civil registration, a legal requirement.

- The statistics are based only on live births, as there is no public register of stillbirths.

- The statistics include live births that occurred in the calendar year but include a very small number of late registrations from the previous year.

- Babies born in England and Wales to women whose usual residence is outside England and Wales are included in the statistics for England and Wales as a whole, but excluded from any subdivision of England and Wales.

- The statistics are based on the exact spelling of the name given on the birth certificate; grouping names with similar pronunciation would change the rankings – exact names are given so users can group if they wish.

4 . Quality summary

Overview of the output

Our baby names release presents data on the first names of live-born babies. The statistics are published annually and represent births occurring in England and Wales. Figures are derived from names recorded when a birth is registered in England and Wales. The release provides counts and ranks for:

- names in England and Wales for baby boys and girls

- the top 100 names in England

- the top 100 names in Wales

- the top 10 names by month of birth

- the top 10 names by mother's region of usual residence

- the most popular name by mother's local authority of usual residence

- the top 100 names by age of mother

The Office for National Statistics (ONS) took responsibility for producing baby name statistics for England and Wales in 2009. Prior to this, figures were produced by the General Register Office (GRO). A time series of baby names for boys and girls back to 1996 is available, providing counts and ranks of names in England and Wales and the top 10 names by month of birth.

For years prior to 1996, the top 100 rankings in England and Wales were produced by the GRO for the years 1904 to 1994 at 10-yearly intervals. These statistics are published on our website. We are unable to provide any more detailed data beyond what is already available for these years, including counts.

For information on data quality, legislation and procedures relating to birth statistics, please see our User guide to birth statistics.

5 . Quality characteristics of the baby names data

(The degree to which statistical outputs meet user needs.)

Our annual baby names release provides statistics on the first names given to live-born babies born in England and Wales in a given calendar year, but includes a very small number of late registrations from the previous year. The release is published around nine months after the end of the data year. A statistical bulletin provides commentary on the published datasets.

There is no publicly available register of stillbirths, therefore baby name statistics are based only on live births.

Baby name statistics do not include births to women usually resident in England or Wales who give birth abroad. They do include births that occurred in England or Wales to women whose usual residence is outside England and Wales. Such births are included in total figures for England and Wales but excluded from any subdivision of England and Wales.

An interactive data visualisation tool has been produced by a web developer external to Office for National Statistics (ONS) and enables users to visually compare changes in the popularity of different names since 1996.

The primary users of the data are parents, soon-to-be parents, and the media. Baby name websites and those who manufacture and sell named items such as souvenir mugs also make use of the data.

The baby name statistics are based on the exact spelling of the name given on the birth certificate. Some users request that similar names be grouped. We provide only statistics based on the exact spelling and do not group names, as some groupings are subjective and not straightforward. Users can create their own groupings if they wish.

It is necessary to protect the confidentiality of uncommon baby names to prevent the identification of individuals and the potential linkage of these data to other datasets. As such, we do not release names with counts of fewer than three in England and Wales as a whole. Further information on the ONS policy on protecting the confidentiality in tables of birth and death statistics is available.

The assessment of user needs and perceptions section provides further information about processes for finding out about uses and users, and their views on the baby names release.

Accuracy

(The degree of closeness between an estimate and the true value.)

Baby names data are based on actual birth registrations. These data represent the legal record, making it the best and most complete data source. As part of the birth registration process, before data are submitted through the Registration Online System (RON), the registrar asks the informant (typically one or both parents) to verify that all data entered are accurate. The registrar is then able to correct any errors. The name supplied will be the name on the birth certificate required in order to obtain a passport or a school place.

The births annual dataset used to produce the statistics is a static file of birth registration records available at the time the dataset is closed. Revisions to records can still be made after the dataset has been finalised, but these will not be reflected in the annual dataset or in published statistics.

Between 1996 and 2000, the cut-off date for inclusion in the annual dataset was births occurring in the reference year that were registered by 11 February of the following year; 42 days after 31 December, the legal time limit for registering a birth. For 2001, the cut-off date was extended to 25 February 2002 to allow increased capture of births registered late. This change means that the annual statistics are prepared on as close to a true occurrences basis as possible without further delay to publication.

Prior to 2001 the annual dataset included:

- births occurring in the reference year that were registered by 11 February the following year

- births occurring in the year prior to the reference year that were registered between 12 February in the reference year and 11 February the following year; that is, births in the previous year that had not already been tabulated

Annual datasets for 1996 to 1999 were derived in a similar way, except that late registrations for births for all earlier years were included in the annual total, not just late registrations for births in the previous year.

Since 2001 the annual dataset includes:

- births occurring in the reference year that were registered by 25 February the following year (parents are allowed 42 days to register a birth)

- births occurring in the year before the reference year but were registered late (after the previous 25 February cut-off) and therefore did not make it into the previous year's statistics

The compilation of baby name statistics has been automated as much as possible to ensure accuracy and efficiency but maintain timely releases. Minimal automated editing is conducted on the names. Part of the automation process involves removing accents from names. Names are analysed without accents since they cannot be automatically processed within the software used.

Prior to the 2018 data year, spaces and any text following a space was removed. Therefore, for names which included any spaces before or after a hyphen, the name following the hyphen would have been treated as a second name. For example, names such as Amelia - Lily would be processed as Amelia. For 2018 data onwards, the process in cleaning names has changed. Now, all spaces are removed from hyphenated names, to ensure they are treated as one name For example, Amelia - Lily would now be processed as Amelia-Lily.

When analysing 2018 data using the old and new cleaning method, the impact of the changes on rankings of baby names has been negligible; none of the top 100 names changed by more than one rank in either direction. The change resulted in a very small increase in the number of unique names in 2018 (129 additional boys' names and 172 additional girls' names), compared with using the old cleaning method.

Coherence and comparability

(Coherence is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar. Comparability is the degree to which data can be compared over time and domain, for example, geographic level.)

We provide a time series of counts and ranks of baby names for boys and girls back to 1996. For years where we are unable to provide detailed data (prior to 1996), the top 100 rankings put together by the General Register Office (GRO) are published for all possible years (1904 to 1994 at 10-yearly intervals). Counts are not available before 1996, which affects the comparability of baby name statistics prior to this date.

Prior to 2009 when the GRO produced statistics on baby names, figures were published several months earlier as they were based only on births registered in the first 46 weeks of the year. This could have introduced some seasonal bias. For example, Holly is a very popular girls name in December and a large proportion of girls called Holly born in the data year would consequently have been excluded from published figures.

The published counts are based on the exact spelling of the first name given on the birth certificate. This is consistent internationally with countries such as Scotland, Northern Ireland, the Netherlands, the US, Canada and the Republic of Ireland. There are, however, some differences internationally. For example, New Zealand uses date of registration rather than date of birth.

We publish our baby names release around nine months after the end of the data year. Baby names for Scotland and Northern Ireland are published by the National Records of Scotland (NRS) and the Northern Ireland Statistics and Research Agency (NISRA) respectively. The ONS, NISRA and NRS all produce baby name statistics using information collected at birth registration for live births only.

We are not the only organisation to produce annual baby name statistics for England and Wales. Bounty (a parenting organisation) produces statistics using voluntary responses received from new mothers. These statistics are not as complete as those produced by us, since not all women giving birth volunteer the information to Bounty.

Concepts and definitions

(Concepts and definitions describe the legislation governing the output and a description of the classifications used in the output.)

Baby names statistics are derived from names recorded when a birth is registered in England and Wales. Birth registration is a legal requirement under the Births and Deaths Registration Act 1836. The registration of births occurring in England and Wales is a service carried out by the Local Registration Service in partnership with the General Register Office (GRO).

Accessibility and clarity

(Accessibility is the ease with which users can access the data, also reflecting the format in which the data are available, and the availability of supporting information. Clarity refers to the quality and sufficiency of the release details, illustrations and accompanying advice.)

Our recommended format for accessible content is a combination of HTML web pages for narrative, and charts and graphs; data is provided in usable formats such as CSV and Excel. Our website also offers users the option to download the narrative in PDF format. In some instances, other software may be used or may be available on request. Available formats for content published on our website but not produced by us, or referenced on our website but stored elsewhere, may vary. For further information please contact us via email at Health.Data@ons.gov.uk.

Special extracts and tabulations of baby names for England and Wales are available to order (subject to legal frameworks, disclosure control, resources and our charging policy, where appropriate). Such enquiries should be made via email to Health.Data@ons.gov.uk or by telephone on +44 1329 444110. We also publish user requested data.

We welcome feedback on the content, format and relevance of releases. Please send feedback to Health.Data@ons.gov.uk.

Baby names for 1996 onwards are available in annual files: one for boys and one for girls. Alongside these, a statistical bulletin provides supporting commentary. The bulletin outlines main findings and describes recent trends.

To aid visual interpretation further there is a baby names interactive data visualisation tool. The tool shows how names have changed in popularity since 1996.

To help users identify changes within the published tables, increases in rank are denoted by a + and decreases with a -. 'New entries into the top 100, when ranks for the year are compared with ranks for another year, are indicated in adjacent columns. A [x] is used to denote when a name has not previously been included in the rankings, that is, it had a count of fewer than three in the comparison year.

Following guidance from the Government Statistical Service (GSS) to improve digital accessibility of statistical spreadsheets, our published datasets for baby name statistics from the 2021 data year now incorporate and adhere to this guidance. This is to help improve the usability, accessibility, and machine readability of our statistical spreadsheets.

Timeliness and punctuality

(Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the gap between planned and actual publication dates.)

The annual release of our baby names bulletin is announced on the GOV.UK release calendar at least four weeks before publication.

Our baby names release is published around nine months after the end of the data year following the full quality assurance of the data. This time lag is necessary to ensure that the statistics are based on the annual births dataset ensuring the highest possible quality.

Our baby names release for 2009 was delayed because of methodological changes to the way the dataset was created. These changes were necessary to enable us to answer customer requests regarding the number of babies registered without a name.

For more details on related releases, the GOV.UK release calendar is available online and provides 12 months' advance notice of release dates. In the unlikely event of a change to the preannounced release schedule, public attention will be drawn to the change and the reasons for the change will be explained fully at the same time, as set out in the Code of Practice for Statistics.

6 . Methods used to produce the baby names data

How the output is created

Baby name statistics are derived from final annual births registration data. They represent all live births occurring in England and Wales in the specific calendar year and include a very small number of late registrations from the previous year. The statistics are based on the exact spelling of the name given on the birth certificate. However, some data cleaning is conducted to ensure names are accurately processed. The compilation of these statistics has been automated as much as possible to ensure efficiency and accuracy but maintain timely releases.

7 . Other information

Assessment of user needs and perceptions

(The processes for finding out about uses and users, and their views on the statistical products.)

A feedback survey for baby name statistics took place in July 2011. The results and responses to this survey were published in August 2012.

We welcome feedback on the content, format, and relevance of releases. Please send feedback to Health. Data@ons.gov.uk.

User feedback is requested at the bottom of all emails sent by customer service teams.

Useful links

Annual baby names are published by month of birth and country and region of usual residence of the mother. An interactive data visualisation tool enables users to visually compare changes in the popularity of different names since 1996.

Our baby names explorer enables users to compare changes in the top 100 baby names for boys and girls since 1904.

Baby names for Scotland and Northern Ireland are published by the National Records of Scotland and the Northern Ireland Statistics and Research Agency respectively.

For information on data quality, legislation and procedures relating to birth statistics, please see our User guide to birth statistics.

8 . Cite This Release

Office for National Statistics (ONS), released 5 October 2022, ONS website, QMI, Baby Names QMI