

HUMAN FERTILITY DATABASE DOCUMENTATION: PORTUGAL

Authors:

Maria Filomena Mendes
Universidade de Évora
E-mail: mmendes@uevora.pt

Isabel Oliveira
ISCTE – Instituto Universitário de Lisboa
E-mail: isabel.oliveira@iscte.pt

Maria da Graça Magalhães
Instituto Nacional de Estatística
E-mail: mgraca.magalhaes@ine.pt

Filipe Ribeiro
Universidade de Évora

Lídia Tomé
Universidade de Évora

Kryštof Zeman
Vienna Institute of Demography, Austrian Academy of Sciences
Email: krystof.zeman@oeaw.ac.at

Last revision: 17 January 2025

WARNINGS:

Due to high migration in the 1960s through the 1980s, the quality of inter-censal population estimates for this period is limited. Cohort fertility indicators for cohorts born prior to 1976 and period fertility indicators for periods prior to 1959 should be used only with special caution. Details are provided in sections 3.2 and 4.5.

1 General information

This report documents the data for Portugal included in the Human Fertility Database Project.

Established in 1935, Statistics Portugal (*Instituto Nacional de Estatística* – INE, IP) is the official Portuguese statistical office. Prior to this date, there were other official institutions responsible for maintaining statistics from the end of the 19th century onwards. Statistics Portugal is responsible for vital statistics collection, for the dissemination of demographic statistics for Portugal, as well as for conducting and preparing the population census.

1.1 Territorial coverage

The Portuguese mainland has had stable geographical boundaries since 1143. In the 16th century, two territories in the Atlantic Ocean, the islands of Madeira and Azores (*Açores*), were added, and have been part of Portugal ever since.

The data collected for the Human Fertility Database consider the actual Portuguese territory that includes both the mainland (*Continente*) and the two insular areas (*Região Autónoma da Madeira e Região Autónoma dos Açores*). The former colonies in Africa and Asia are not included in these figures.

1.2 Data collection and availability

Vital statistics were collected by a number of different official institutions until 1935. Since then, data on births, deaths, and marriages have been produced and published by Statistics Portugal.

Population censuses have been carried out regularly since the last quarter of the 19th century. The population censuses were conducted in 1864, 1878, 1890, 1900, 1911, 1920, 1930, 1940, 1950, 1960, 1970, 1981, 1991, 2001, 2011, and 2021. Since 1940, the censuses have been organized and published by Statistics Portugal. The 1970 census was not completed, so the data for the 1970 population are based on a 20% sample from the officially published estimates.

Data collected for the HFD consist of age-specific live births for 1930–1954¹, age- and birth order-specific live births for 1955–2023, census data on the parity distribution of women by age in 1981 and 1991, and monthly numbers of live births since 1901.

Population data since 1940 have been processed and documented in the Human Mortality Database (HMD, www.mortality.org).

All the input data used for generating the HFD output data and indicators are specified in Appendix 1.

1.3 The problem of recording migration

Substantial migration movements took place during the whole 20th century, but these flows became particularly significant in the 1960s and affected the reconstruction of population numbers and also the indicators of cohort fertility. In 1960, the Portuguese population was about 8.8 million; in 1970 it was 8.7 million. During this decade the net migration was negative and approached 1.3 million, consisting mainly of people in fertile ages.

During the 1950s, and even more so in the 1960s, migration to the former Portuguese colonies in Africa was widespread. However, in the next decade, after the 1974 revolution, Portugal lost the African colonies and about 500 to 700 thousand people returned from the newly independent African countries. These immigrants included a large number of females in fertile ages. The return had occurred very rapidly between the second half of 1974 and 1975, also because of the civil war in Angola and Mozambique.

In the 1980s migration was comparatively mild. In the 1990s Portugal became a country of immigration, particularly from Latin America and Eastern Europe.

¹ Since the population estimates in the HMD are available only since 1940, the HFD does not compute fertility indicators for the period 1930–1939. Nevertheless, the birth counts for this period are available in the input file.

2 Birth count data

Compulsory registration of births began in 1911. Prior to this date, records of baptism were kept by the local clergy.

Data on births released by Statistics Portugal are based on the information provided by individuals at the time of the administrative record procedure at the Civil Registration Offices.

The data on births since 1980 have been provided by Statistics Portugal on request, and comprise the following information on live births of mothers resident in Portugal: annual live births by birth order, single years of age, and birth cohort of the mother; live births by year and month of birth. Prior to this date, the data file on births was based on the tabulation data from the demographic yearbook² available in the Digital Library of Statistics Portugal (www.ine.pt).

These historical birth data are available with varying levels of detail:

- a) The total number of live births by sex has been known since 1886, but only after 1930 a more detailed distribution of births became available.
- b) For the years 1930 to 1954, data on live births are available by five-year age groups of mothers only, including the broad age group –19. Because such a broad age group cannot be satisfactorily recalculated into single age groups using current HFD methodology (see the HFD methods protocol), data are first recalculated into narrower age groups 12–14 and 15–19 using the method described in Appendix 2.
- c) For the years 1955 and 1956, data on live births by age of mother (in single years of age) and birth order are available for children born within marriage only; for children born outside wedlock (about 11% of all live births), the age of mother and birth order are unknown. This may pose a problem because the age structure of births to non-married women is then estimated from the age structure of marital births, which is considerably different: the mean age at birth among non-married mothers is by about 2.5 years lower and their age schedule of childbearing is shifted to younger ages when compared with the married mothers (according to data from 1957, see Figure 1).
- d) For the years 1957 and 1958, data on live births are specified by age of mother and birth order; for children born outside marriage, the birth order is unknown. This may pose a problem because the parity structure of non-marital births is different from that of marital births. Non-marital births are mostly first or second births, so when the unknown cases are recalculated according to the parity structure of the marital births, the resulting parity structure will underestimate the number of first births (see Figure 2).

Therefore, birth order specific data for 1955–1958 are not used for the computation of HFD output.

- e) For the years 1959 to 1979, data on live births by age of mother (single years of age until 50 and over) and birth order are available;
- f) For the years 1970 and 1971, the numbers of births by age of mother were re-estimated by Statistics Portugal, based on the initially reported relative age distribution (released in *Estatísticas Demográficas*, 1970 and *Estatísticas Demográficas*, 1971) and subsequently revised absolute numbers of total births.

² Through 1967, 38 volumes with the title *Anuário Demográfico* were published. From 1968 onwards, the yearbook was renamed *Estatísticas Demográficas*. However, in the digital library all these materials are available under the heading *Estatísticas Demográficas*.

Figure 1 Proportion of live births by age in 1955–1957, married and non-married mothers

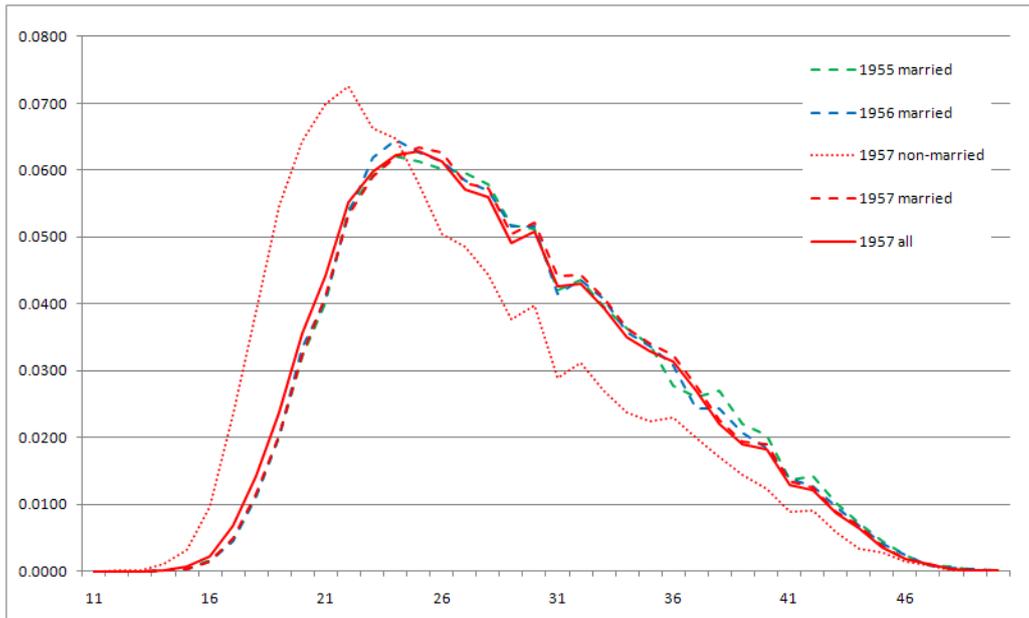
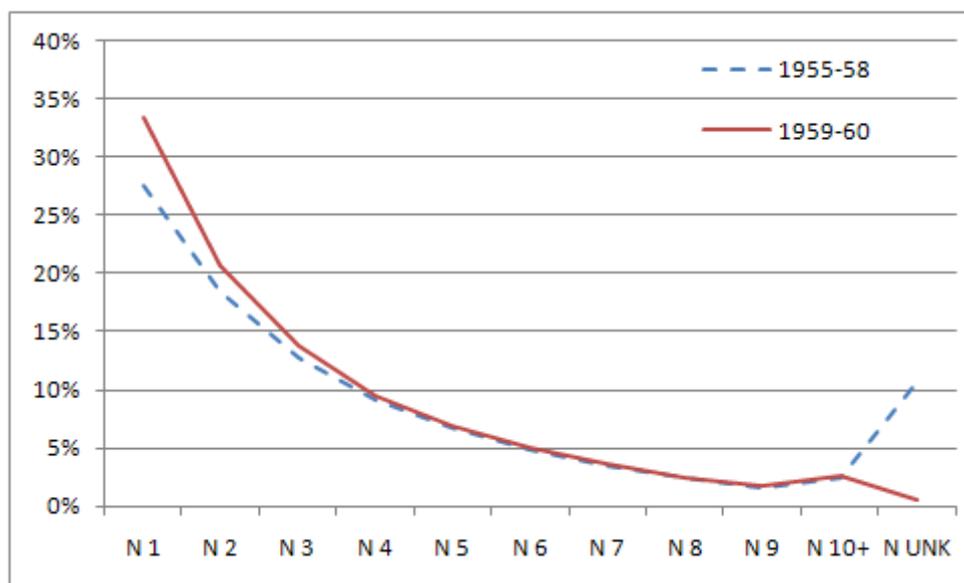


Figure 2 Proportion of live births by birth order in 1955–1960; for 1955–1958, birth order is unknown for all non-married mothers



3 Population count data

3.1 Population count data by age

The massive and changing migration in the 1960s–1990s, and the lack of proper census in 1970, makes it very difficult to make accurate population estimates. Following the standard HFD approach, the annual age structure of women as at January 1 for the years since 1940 is taken from the Human Mortality Database (HMD). It is noteworthy that the data from the HMD considerably differ from official data of Statistics Portugal (see Canudas-Romo et al. 2023).

Statistics Portugal releases annually estimates for resident population by single years of age (up to 85 and over) and sex. These population estimates are of two types: post-censal, based on the most recent population census combined with vital statistics data for the subsequent years, and therefore considered "provisional"; and inter-censal ones, that revise the former estimates and are based on the results of two consecutive censuses. The post-censal annual population estimates are computed by the component method, applied for each sex separately. They combine resident population counted at the last census with administrative records of live births (of resident mothers) and deaths (of residents) by sex and year of birth as well as estimated migratory flows by sex and single years of age (based on the administrative data from the Foreign Nationals and Borders Service (SEF) and on indirect sources such as the Labour Force Survey (LFS) or a module on "Emigration Survey" from the LFS).

The HMD population estimates proportionally distributes migration over the inter-censal period and does not take the year-to-year migration fluctuations into account (for more details, see the HMD Methods Protocol), while the statistical office considers its estimated short-term trends. Both estimates apparently have their advantages and flaws. The HMD approach was applied for 1940–1990. For the three inter-censal periods 1991–2000, 2001–2010, and 2011–2020, the HMD uses the INE inter-censal estimates for the input data. The population estimates for 2011–2020 were recalculated, considering the 2011 and 2021 census data (Canudas-Romo et al. 2024). The population estimates for 2021 are considered final. Population estimates for 2022 were revised in the last HMD release. Differences in the population estimated by the HMD approach and the approach employed by Statistics Portugal (see Figure 3) explain the gap between the official TFRs and the TFRs computed by the HFD in the order of 0.02–0.12 in 1961–1968 and 1984–1987 (see Figure 4).

Figure 3 Number of women at ages 20–34 according to the official estimates and those of the Human Mortality Database, 1940–2023

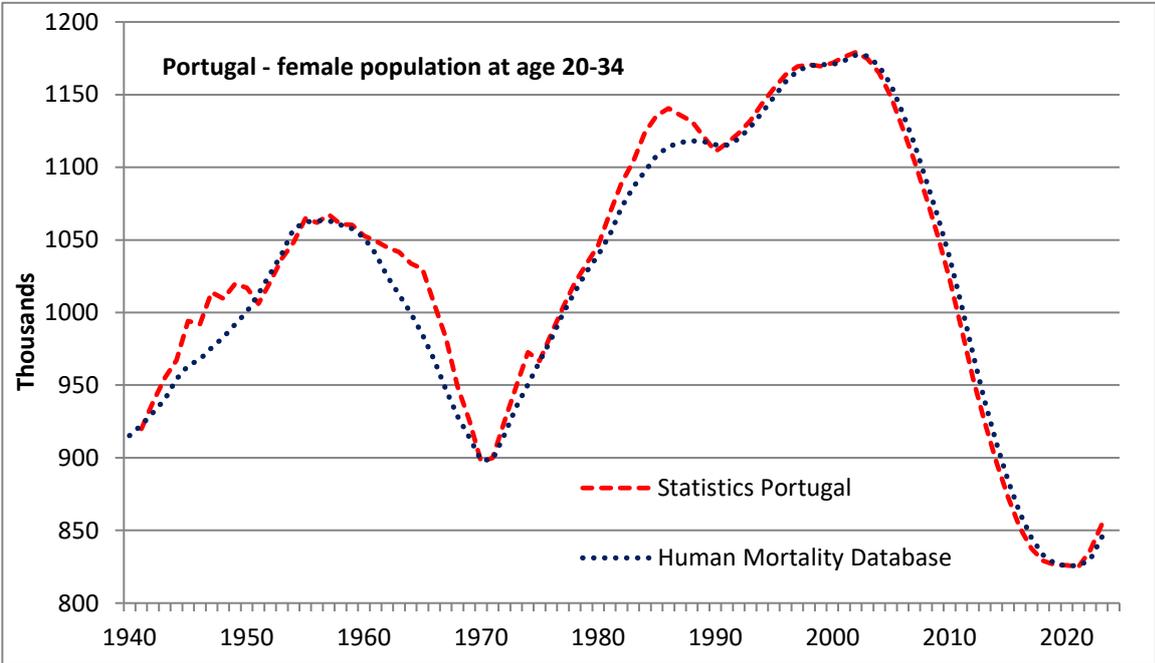
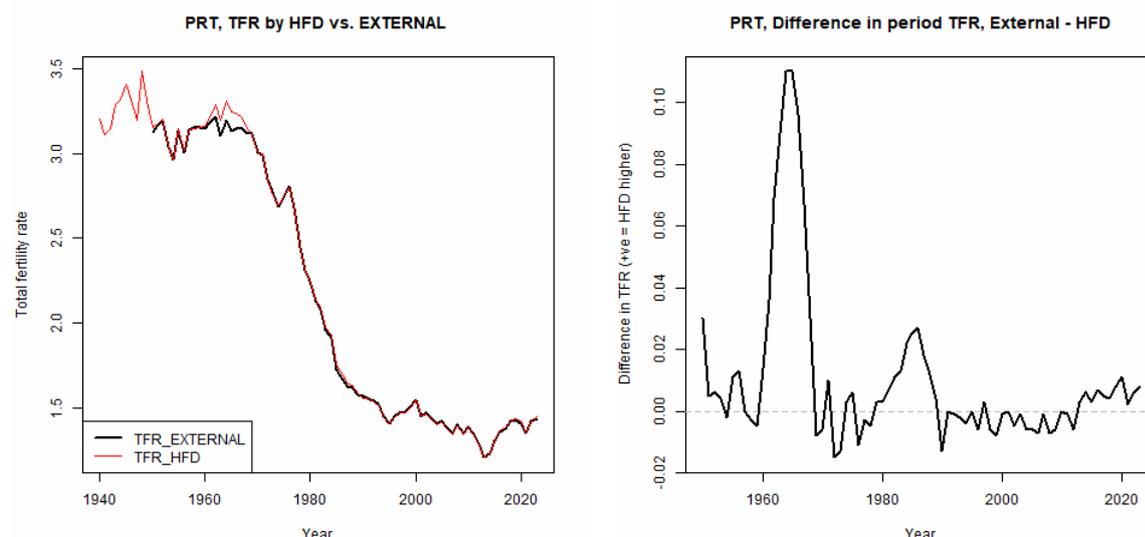


Figure 4 Comparison of the official TFRs and those calculated by the HFD, 1940–2023



Note: Official estimates, presented as “external” in the graph, come from two sources: Eurostat (for 1950–1970) and the official website of Statistics Portugal (since 1971).

3.2 Population count data by age and parity

Data on women by age and the number of live-born children are available from the censuses of 1981 (March 16) and 1991 (April 15). For the previous periods, the female population by birth order is available only from some censuses, and only for five-year age groups; this information is, however, insufficient to be used for HFD purposes³. In the censuses of 2001, 2011, and 2021, the question on parity was not included.

The parity composition of women differs substantially when estimated using different sources: the censuses of 1981 and 1991 as well as the cumulation of cohort age-specific fertility rates give contrasting results. As shown in Figure 5 and Figure 6, while the cumulation of vital statistics data underestimates the proportion of childless women, which reaches implausible low levels, it also overestimates the proportion of women with one child. This mismatch is more evident for older women. We attribute this difference to the high migration in inter-censal periods.

Under regular circumstances, given the available data, period fertility tables would be calculated using age- and parity-specific female population weights generated by cumulating fertility rates over the reproductive lifespan of cohorts. However, due to the aforementioned reasons, we use the ‘golden census’ approach where the female parity distribution by age from the 1991 census is employed as a starting point for the computation of period fertility tables (for more details on the two approaches, see the HFD Method Protocol).

In any case, all the cohort fertility indicators computed in the HFD should be used with caution, especially those of cohorts born before 1976 (i.e. the year of the census 1991 minus 15 years of age).

³ The census collected information on women’s parity since 1940 by five-year age groups. In 1940 it considered only the married women, and the data were inaccurate. In 1950 the census data included only currently married women; and in 1960, only ever-married women. The 1970 census was not completed, so the data for the 1970 population are based on a 20% sample from the officially published estimates, and the data on the number of children contain gross errors.

Figure 5 Childlessness by age in 1991, estimated using different sources

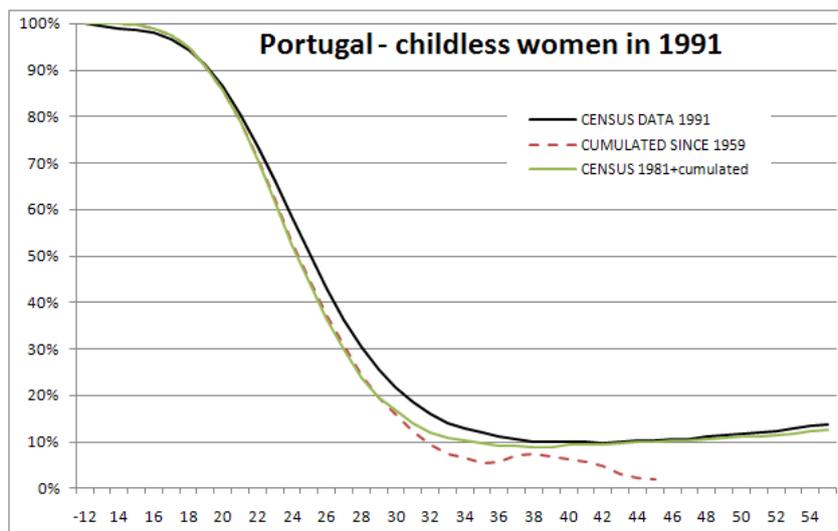
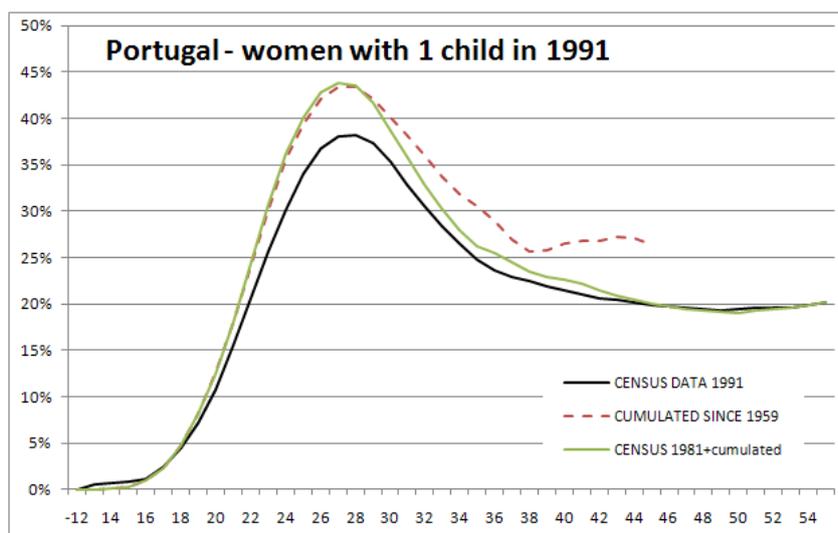


Figure 6 Proportion of women with one child by age in 1991, estimated using different sources



4 Specific details

4.1 Definitions of live birth

From 1968 onwards, each demographic yearbook, or *Estatísticas Demográficas*, featured a section in which concepts were defined (*Conceitos e Anotações*), including that of live birth. Since then, no major changes have been made, and the definition of live birth complies with the WHO definition: “Live birth is the complete expulsion or extraction from the mother’s body, irrespective of the duration of the pregnancy, of product of conception that, after separation, breathes or shows any other signs of life, such as a beating of the heart or umbilical cord or actual contraction of any muscle subject to voluntary movement, whether or not the umbilical cord has been cut and whether or not the placenta is retained.”

4.2 Age

The birth data for 1930–1979 are tabulated by age in completed years (Lexis squares).

Starting in 1980 the data are available by both mother's age in completed years and mother's birth cohort, recalculated into age reached during the year. The Statistics Portugal definition of age is "the interval of time between the day, month and year of birth and 0:00 a.m. of the reference date. Age is expressed in completed years for adults and children and in months, weeks, days, hours or minutes of life, as appropriate, for infants under one year of age".

4.3 Birth order

The live birth order is defined as the number of previous live births of the woman, plus the birth which has just occurred.

4.4 Unknown cases

In 1930–1954 birth order was not recorded, and the proportion of births with unknown age was around 1%. For 1955–1956 age of mother and birth order are unknown for children born outside wedlock (around 11% of all births). For 1957–1958 birth order was unknown for children born outside wedlock (around 11% of all births). The proportion of unknown cases (either of unknown age, or birth order, or both) in 1959–1971 was decreasing from 0.6% to 0.1% and thereafter it stayed below 0.1%. In 2006, 2007, 2009 and 2011–2012 the proportion of births with unknown birth order was 0.1–0.2%, while in 2008 both age of mother and birth order were unknown for 1% of births. Since 2013 there are no unknown cases; in 2019 there was one case where unknown age and birth order of the birth were unknown, in 2021 there were two cases with unknown age of mother, and in 2023 there was one case with unknown birth order.

4.5 Quality of the data

The quality of population data is affected by high migration in the intercensal periods (see Section 3.1). Due to the problems in matching births with the female population, and lower quality of the data prior to 1959, cohort fertility data derived using the standard HFD methodology of cumulating age-specific fertility rates of cohorts observed over long periods of time may reach implausible results (for instance, childlessness of cohorts 1944–1948 is lower than 5% and fluctuating – see also Section 3.2). Due to the aforementioned reasons, we use the 'golden census' approach where the female parity distribution by age from the 1991 census is employed as a starting point for the computation of the period fertility tables (for more details, see the HFD Method Protocol).

All the cohort fertility indicators computed in the HFD should be used with caution, especially those of cohorts born before 1976 (i.e. those reaching age 15 or higher at the 1991 population census).

Also period fertility indicators for the period prior to 1959 should be used with caution, due to several quality concerns, including the signs of age heaping prior to year 1955, and redistribution of unknown age for non-married mothers in 1955–1956.

4.6 Revision history

Changes with the March 2014 revision:

The data release employs new population estimates for 1991–2010 based on the 2011 census results. Other changes include modified population estimates for 1981–1990 (for details, see the HMD Documentation for Portugal by Canudas-Romo et al. 2024). Therefore, in this data release fertility rates and other indicators for 1981–2009 differ from those released on November 1, 2011, but at the same time they have become closer to the official figures published by Statistics Portugal.

Changes with the March 2017 revision:

The release includes new data for 2013–2015.

Changes with the April 2020 revision:

The release includes new data for 2016–2018.

Changes with the August 2021 revision:

The release includes new data for 2019–2020.

Changes with the September 2022 revision:

The release includes new data for 2021. Numbers of births for 2020 were updated, 104 births have been added by Statistics Portugal, causing negligible changes in related indicators.

Changes with the February 2024 revision:

The release includes new data for 2022. Population estimates for 2011–2021 were re-estimated, considering the 2011 and 2021 census data (Canudas-Romo et al. 2024). Therefore, there were changes in the age-specific fertility rates and also in the values of the summary indicators for this period. The total fertility rate increased by 0.002–0.015, and the largest increase was in 2020 (from 1.406 to 1.421).

Changes with the January 2025 revision:

The release includes new data for 2023. Population estimates for 2022 were updated (Canudas-Romo et al. 2024). Therefore, there were changes in the age-specific fertility rates and in the values of the summary indicators for this period. The total fertility rate estimated for 2022 changed from 1.435 to 1.426.

Acknowledgements

We would like to thank Statistics Portugal for their cooperation in the collection of data.

References

Canudas-Romo, Vladimir, Edviges Coelho, Cláudia Antunes Pina, Gabriel Borges, Timothy Riffe, Felipe Menares, Denys Dukhovnov, Emma Lonstrup. 2024. *About mortality data for Portugal*. Documentation report for the Human Mortality Database. Last updated 4.11.2024. Available online on www.mortality.org

APPENDIX 1 INPUT DATA USED FOR HFD CALCULATIONS

BIRTHS

Period	Type of data	Age range	Birth order	RefCode(s)	Notes
1940–1954	Annual number of live births by age of mother (Lexis squares)	–19, 20–24...45–49, 50+, unknown	—	1	
1955–1956	Annual number of live births by age of mother (Lexis squares)	11, 12, ..., 49, 50+, unknown	—	1	unknown age and birth order for non-married mothers; order-specific data therefore not used
1957–1958	Annual number of live births by age of mother (Lexis squares)	11, 12, ..., 49, 50+, unknown	—	1	unknown birth order for non-married mothers; order-specific data therefore not used
1959–1979	Annual number of live births by age of mother and birth order (Lexis squares)	11, 12, ..., 49, 50+, unknown	1, ..., 10+, unknown	1, 2	
1980–2023	Annual number of live births by age of mother, mother's year of birth and birth order (Lexis triangles)	11,...,61, unknown	1, ..., 10+, unknown	3, 6, 7, 9, 10, 11, 12	
1901–1910, 1918–1925, 1929–2023	Annual number of live births by month	—	—	1, 2, 3, 6, 7, 9, 10, 11, 12	

FEMALE POPULATION: Distribution by age and parity

Period	Type of data	Age range	Parity	RefCode	Notes
16.3.1981	Number of women by age and parity	0,...,100+	1, ..., 9+	4	no unknown parity
15.4.1991	Number of women by age and parity	0,...,100+	1, ..., 9+	5	'Golden census'; no unknown parity ⁴

FEMALE POPULATION: Exposure by age and year of birth

The female exposure population by calendar year, age, and year of birth (Lexis triangles) is estimated using data on population size and deaths from the Human Mortality Database, which are available at <http://www.mortality.org> or <http://www.humanmortality.de>.

⁴ For Portugal, a country-specific method for the computation of period fertility tables is used. Instead of cumulating age-specific fertility rates to get annual age-parity distributions of women, which is usually done in such cases, we apply the 'golden census' approach (for more details on the two approaches, see the Methods Protocol).

APPENDIX 2

ADJUSTMENTS OF BIRTH DATA BELOW AGE 20

For the years 1930–1954, birth counts by age at young ages are available only for a broad age category –19. Because such a broad age group cannot be satisfactorily recalculated into single age groups using current HFD methodology (see HFD methods protocol), data are first recalculated into narrower age groups 12–14 and 15–19 using the following simple method:

First the ratio of births in the narrower age group 12–14 to births in the broader age group 12–19 is estimated using the known figures from the years 1957–1960⁵:

$$r(12-14) = \frac{B(\leq 14, 1957-1960)}{B(\leq 19, 1957-1960)} \quad (\text{A1})$$

Empirically, in the case of Portugal, $r(12-14) \cong 0.33\%$. Then the number of live births at age 12–14 for each year 1930–1954 is estimated as:

$$B(12-14, t) = r(12-14) \cdot B(\leq 19, t) \quad (\text{A2})$$

The remainder is the number of live births at age group 15–19:

$$B(15-19, t) = B(\leq 19, t) - B(12-14, t) \quad (\text{A3})$$

⁵ Data for 1955–1956 are not used because births to non-married women were added into the category with unknown age of mother in this period.