

### KEY FINDINGS

- There are five main sources for population figures in Ukraine. These all use the 2001 census as a baseline but vary in the methods used to estimate the country's current population. As a result, discrepancies have existed between population estimates and actual figures even before the Russian invasion in February 2022. The last census in Ukraine was conducted in 2001, and plans for a new census have been postponed several times (Hromadske 21/04/2020).
- IOM has two figures on the number of IDPs in the country. Its General Population Survey disaggregates the figure at the oblast level and provides estimates based on phone surveys (IOM 03/07/2023). The humanitarian response in Ukraine usually uses this figure. On the other hand, the Area-Based Assessment features a greater disaggregation that reaches down to the raion level but only counts IDPs officially registered with the Government of Ukraine, resulting in an undercount (IOM 14/07/2023).
- The University of Oxford has estimates based on Facebook user counts of the total population after February 2022 (Leasure et al. 11/05/2022). Their data excludes Crimea and Sevastopol. The figures for Luhansk and Donetsk are likely unreliable given the lack of reported Facebook activity in these oblasts.
- The World Bank's 2022 population figure is also likely to be a post-invasion estimate based on the difference between the 2021 and 2022 estimates, but the data is only available at the national level, and we were unable to find information on what period of the year it covered (WB accessed 12/07/2023).
- We did not find any fundamental incompatibilities between the datasets, likely because the open sharing of data allowed for comparing and refining estimates and because total population counts were based on the same baseline (the 2001 census).
- UNFPA produces the Common Operational Dataset on Population Statistics (COD-PS), which has the only pre-invasion population estimate that included Crimea and Sevastopol (UNFPA 14/11/2023).
- Four sources of population data look at a closer level than the oblast level, including the raion and hromada levels, all with certain limitations. The State Statistics Service of Ukraine (SSSU) figures, and Worldpop's breakdown of UNFPA's estimates are both pre-invasion

population estimates, and Worldpop's dataset requires geographic information system (GIS) software and skills; IOM's raion-level IDP data, only covers officially registered IDPs in government-controlled areas; and the University of Oxford's updated population data at the hromada level, is not publicly available (SSSU 31/05/2022; WorldPop accessed 07/03/2023; IOM 14/07/2023). These limitations create difficulties for humanitarian organisations who want to create targeted response programmes (expert interview 22/05/2023).

- Only IOM's IDP data contains information on people with disabilities (IOM 03/07/2023; IOM 14/07/2023).
- The only pre-invasion dataset that includes all of the currently non-government controlled areas (NGCAs), Crimea, and Sevastopol is UNFPA's COD-PS dataset (UNFPA 14/11/2023).

### About this report

This report explores and explains their methodologies, strengths, and limitations of the total population, refugee, IDP, and returnee datasets commonly used in the Ukraine humanitarian response. It describes population figures from data sources, academic institutions, and international organisations, such as the University of Oxford, IOM, UNHCR, OCHA, and the World Bank.

**Aim:** this report seeks to provide humanitarian responders and stakeholders with an understanding of the characteristics, including the strengths and limitations, of each source on population figures. The aim is to facilitate using the most relevant databases in humanitarian analysis and response planning.

**Method:** the information in this report was gathered via a secondary data analysis of publicly available information on various forms of population data and through discussions with key experts involved in some of the datasets described.

**Limitations:** this report was based on publicly available information as at 19 July 2023. Updates to the numbers in the datasets render the report's numbers outdated. The critical analysis remains relevant until the methodologies of different datasets change.

## TABLE OF CONTENTS

Using data on Ukraine.....	2
Baseline total population figures.....	3
Post-invasion total population figures.....	5
IDPs and returnees.....	6

## USING DATA ON UKRAINE

Data is necessary for an effective humanitarian response. Accurate population, IDP, and returnee datasets allow humanitarians to preposition aid, distribute lifesaving assistance, calculate the number of people in need, provide concrete figures for advocacy, and conduct monitoring and evaluation to improve programming effectiveness. In an active conflict, where people are frequently moving and data collection on population levels is difficult, accurate data becomes even more important. The selection of data sources depends on the available information, what humanitarian responders need (in terms of geographical scope, timeline, and granularity), and the data's limitations. The more precise and localised a dataset's time and geographic scope, the more assumptions and caveats are needed to produce figures. Precise figures require the collection of data in situations with high access and communication constraints and frequent budget limitations, which can force the use of alternative calculations.

The lack of local-level data in Ukraine makes it difficult for humanitarians to assess the potential impact of displacement and return flows on host communities, because up-to-date information on the size of potential host communities is not available (expert interview 22/05/2023).

**The SSSU conducted the last census in 2001.** Budgetary constraints, the COVID-19 pandemic, and Russia's February 2022 invasion have postponed subsequent attempts (Eurasianet 12/08/2020; SSSU accessed 24/03/2023). **All total population figures since 2001 are based on estimates using different methodologies** and data from before the February 2022 invasion. **Only the University of Oxford and the World Bank provide estimates of post-invasion population numbers.** The University of Oxford does so based on daily usage data

collected via Facebook. The World Bank does not specify that its 2022 figure is post-invasion, but it can be inferred given the sharp drop in its population estimates between 2021–2022 (WB accessed 12/07/2023).

While data disaggregated by age and gender is available in most datasets, data on disability is less easily available. **Only IOM's IDP data contains information on people with disabilities.**

Table 1. Data sources for population figures as at 19 July 2023

DATA	SOURCE	RELEASE DATE	REFERENCE DATE	UPDATE FREQUENCY	SMALLEST GEOGRAPHIC UNIT	GENDER/AGE DISAGGREGATION
Total population (baseline)	State Statistics Service	2022	2021	Yearly	Hromada	Yes
Total population (baseline)	World Bank	2023	2022	Yearly	National	Yes
Total population (baseline)	UNFPA	14/11/2022	01/01/2022	Yearly	Oblast	Yes
Total population (baseline)	WorldPop	11/08/2022	01/01/2022	Unknown	100m2 cells	Yes
Total population (post-invasion)	University of Oxford	26/08/2022 (latest public release)	01/07/2022	Daily	Oblast (latest public release)	Yes
IDPs	University of Oxford	26/08/2022	01/07/2022	Daily	National	Yes
IDPs (registered and unregistered)	IOM	03/07/2023	06/2023	Quarterly	Oblast	Yes
IDPs (registered)	IOM	14/07/2023	06/2023	Monthly	Raion	Yes
Returnees (from abroad and Ukraine)	IOM	03/07/2023	06/2023	Quarterly	Oblast	Yes

Sources: SSSU 31/05/2022; WB (accessed 12/07/2023); UNFPA 14/11/2022; WorldPop (accessed 05/04/2023); Leasure et al. (11/05/2022); Leasure et al. (15/05/2022); IOM (03/07/2023 and 14/07/2023)

**Humanitarian responders looking to plan activities at the local level do not have many options for up to date population data at the raion or hromada level.** SSSU's data provides population estimates at the hromada level as at 1 January 2022. WorldPop's breakdown of pre-February 2022 population figures into a grid of 100m<sup>2</sup> cells allows for the data to be compiled into the desired administrative unit, but retrieving the information from the files requires GIS software and skills. Both these sources do not reflect the impact of displacement since February 2022. The University of Oxford provides current population estimates at the hromada and 2km<sup>2</sup> levels, although those figures are not publicly available. IOM provides a count of registered IDPs at the raion level, but the number of registered IDPs is lower than the total estimated number in the country.

**Responders working at the national level may not need such granularity and may find datasets from IOM, UNHCR, and UNFPA more straightforward.** It is possible to triangulate or adjust monthly or yearly updates on population estimates to the population figures from the University of Oxford. **Those who want additional information on urban/rural breakdown or unemployment rates can use the World Bank's data to complement the population figures.** Pre-invasion population figures from UNFPA or SSSU provide a stable and commonly used baseline and can be updated using UNHCR's data on refugees, as ACAPS has done in its *Ukraine Master Dataset*. WFP's hunger map also uses the World Bank's total population estimates (WFP accessed 04/04/2023).

UNFPA's COD-PS dataset is the only pre-invasion dataset that includes current NGCAs, Crimea, and Sevastopol. Its use of P-codes<sup>1</sup> for oblasts and cities means it is easy to integrate it with OCHA's subnational administrative boundaries and data for people in need (UNFPA 14/11/2023). It is also the most commonly used population baseline in the response. Using P-codes makes it possible to link the data with any other dataset that includes P-code fields for different administrative units. The 2023 Humanitarian Needs Overview for Ukraine uses COD-PS' population figures as the baseline pre-February 2022 total population (OCHA 28/12/2022).

**Data on IDPs is available from IOM and the University of Oxford,** though both datasets have limitations. Most humanitarian analyses use IOM's figures, which are based on representative key informant interviews via telephone. This method means there may be an underrepresentation of residents from areas with significant damage or temporary telecommunication cuts, as they may be challenging to reach (IOM 03/07/2023). The University of Oxford's data can only be used as an estimate, and publicly available data is only available up to June 2022. Because the university relies on Facebook data, it does not provide a complete picture of oblasts where there are both in- and outflows of people, as they may cancel each other out.

<sup>1</sup> P-codes are unique codes that identify a specific administrative unit, from the national to the smallest administrative level (OCHA accessed 23/03/2023).

## BASELINE TOTAL POPULATION FIGURES

Because all these datasets are pre-invasion estimates based on the 2001 census, the following datasets do not consider refugee flows since February 2022.

### State Statistics Service of Ukraine

**Total population estimate as at January 2022:** 41,167,335 (SSSU 31/05/2022).

**Methodology:** the SSSU made their estimates using available administrative data on official birth and death registrations, changes in residence registrations, and migration (SSSU 31/05/2022).

**Strengths:** these official government figures provide a strong pre-invasion baseline estimate down to the hromada level using birth and death registration and international and internal migration data. SSSU's yearly population statistics also contain other interesting demographic indicators, such as birth and death rates, life expectancy, and migration data (SSSU 31/05/2022 and SSSU 21/10/2022).

**Limitations:** as birth and death registrations are based on the place of registration instead of actual residence, there is likely an undercount of people in rural areas who had to travel to urban centres to give birth and registered the births in those urban centres (UNFPA 14/11/2023). This limitation extends to other datasets using SSSU's data for their projections. This group includes UNFPA, whose data was used for the 2023 Humanitarian Needs Overview. The population figures of Donetsk and Luhansk appear to underestimate the number of children under four and overestimate people over 80, as birth and death registrations stopped being communicated to the Ukrainian statistical system in 2014.

**Geographic disaggregation and coverage:** population figures are available at the national, oblast levels, raion, and hromada levels not including the population of Crimea and Sevastopol (SSSU 31/05/2022).

**Age disaggregation:** five-year age groups until 80+ (SSSU accessed 02/06/2023)

**Sex disaggregation:** yes

**Update frequency:** the SSSU updated their figures annually until January 2022. There will be no updates until three months after the end of martial law (SSSU accessed 24/03/2023).

## UNFPA's Common Operational Dataset on Population Statistics

**Total population estimate as at January 2022:** 43,320,154 (UNFPA 14/11/2023)

**Methodology:** UNFPA's COD-PS projections used SSSU's annual birth and death registration data and administrative data on internal and international migration. They adjusted the raw 2001 census data to mitigate an estimated 10.5% average undercount of the oblast population (UNFPA 14/11/2023). The SSSU conducted a pilot census that estimated this undercount in 2019 (UN 01/06/2020).

**Strengths:** it is the most commonly used baseline population figure in the response. It is easily integrated and aligned with OCHA's subnational administrative boundaries and the data for people in need, as it includes P-codes for oblasts and cities (UNFPA 14/11/2023). It can also be linked with any other dataset that includes P-code fields for different administrative units, namely at the national, oblast, raion, and hromada administrative levels.

**Limitations:** this dataset shares the same limitations as SSSU's data, namely that birth and death registration is based on the place of registration, not occurrence (UNFPA 14/11/2023).

**Geographic disaggregation and coverage:** national and oblast data is available. It also provides population estimates for cities with more than 100,000 people.<sup>2</sup> The dataset includes all Ukrainian territory, including pre-February 2022 NGCAs, Crimea, and Sevastopol (UNFPA 14/11/2023).

**Age disaggregation:** five-year age groups until 80+

**Sex disaggregation:** yes

**Update frequency:** yearly

## WorldPop

**Total population estimate as at January 2022:** 43,320,154, the same total population figure as COD-PS' January 2022 figure

**Methodology:** WorldPop estimates the number and density of people living in a given map area, referred to as a cell. Each cell represents 100m<sup>2</sup>. Estimates are produced using UNFPA's COD-PS' January 2022 figures and modelling of built-up areas based on satellite imagery to estimate population density (WorldPop accessed 03/03/2023; WorldPop accessed 07/03/2023). The total population figure equals COD-PS but breaks the population down into 100m<sup>2</sup> cells.

**Strength:** further geographic disaggregation of the most commonly used total population figure in the response (COD-PS).

**Limitations:** this source disaggregates age and sex at the oblast level and applies this uniformly down to the hromada level, ignoring urban-rural differences in age and sex ratios.

**Geographic disaggregation and coverage:** data is available in 100m<sup>2</sup> cells and includes all of Ukraine's territory, including pre-2022 NGCAs, Crimea, and Sevastopol.

**Age disaggregation:** five-year age groups until 80+, with limitations

**Sex disaggregation:** yes, with limitations

**Update frequency:** the time between updates is unknown and likely depends on the release of updated UNFPA COD-PS data.

<sup>2</sup> These cities are Berdiansk, Bila Tserkva, Brovary, Cherkasy, Chernihiv, Chernivtsi, Dnipro, Ivano-Frankivsk, Kamianske, Kharkiv, Kherson, Khmelnytskyi, Kremenchuk, Kropyvnytskyi, Kryvyi Rih, Lutsk, Lviv, Melitopol, Mykolaiv, Nikopol, Odesa, Pavlohrad, Poltava, Rivne, Sumy, Ternopil, Vinnytsia, Uzhhorod, Zaporizhzhia, and Zhytomyr.

## POST-INVASION TOTAL POPULATION FIGURES

### University of Oxford

**Total publicly available population estimate as at July 2022:** 36,325,488 (Leasure et al. 11/05/2022)

**Methodology:** using UNFPA's 2022 COD-PS population figures as a baseline, the University of Oxford combines daily counts of active Facebook users available via Facebook's marketing application programming interface and UNHCR's daily counts of Ukrainian refugees. It then uses daily oblast-wide Facebook user counts to extrapolate population changes and applies these to the 2022 COD-PS baseline. Given the model's reliance on Facebook, data scientists behind the model had to account for the fact that Facebook usage is unequal across age ranges (Leasure et al. 11/05/2022).

**Strengths:** this is the only publicly available dataset with post-invasion total population estimates at the oblast level.

**Limitations:** the data for Luhansk and Donetsk is likely unreliable because the oblasts had low Facebook usage even before the 2022 invasion, and user activity has been at zero since 12 March 2022. Facebook also conducts data-cleaning processes before the user counts are made available, and it is not known how these affect the data (Leasure et al. 11/05/2022).

**Geographic disaggregation and coverage:** the data is available at the national and oblast levels. The University of Oxford's dataset does not include population estimates for Crimea and Sevastopol because Facebook data is unavailable for those locations. More recent data down to the hromada level and below exists but is not publicly available because of the sensitivity of the information. The box below gives the methodology, strengths, and limitations of this more granular data.

**Age disaggregation:** 5-year age groups until 80+. As Facebook does not disaggregate data for users over 65, the University of Oxford disaggregated the data using calculations based on age and sex proportions in the 2022 UNFPA COD-PS. Because of the low user counts of people under 20, proportions from the 2022 UNFPA COD-PS were also adjusted based on the assumption that the number of children remains the same, proportionate to the number of women aged 20–49 (Leasure et al. 11/05/2022).

**Sex disaggregation:** yes

**Update frequency:** daily population estimates until 1 July 2022.

### Updated Oxford total population figures for 2023

The University of Oxford has produced updated population and displacement figures until May 2023. They do not share them publicly because these figures contain population data at the hromada level and below, which is considered sensitive. A public version of the updated methodology and dataset that includes population estimates for the national, oblast, and raion levels will be published soon.

**Methodology:** the researchers applied a machine learning model to their existing data to identify the key factors affecting population change at the oblast level, such as proximity to conflict events and border crossings. The most influential factors are modelled in a grid of 100m<sup>2</sup> cells to produce the predicted proportional population change for each cell, which was then multiplied using WorldPop's baseline population for each cell to produce the current population at a given point in time. The cells are then added up to produce population estimates in a grid of 2km<sup>2</sup> cells and at the hromada level.

**Strengths:** it is the only post-invasion dataset of population figures below the oblast level.

**Limitations:** the updated dataset has the same limitations as the University of Oxford's original dataset. The University of Oxford's researchers have said the figures at the hromada and 2km<sup>2</sup> cell levels from July 2022 onwards should be used with caution as the model performance declines after that date. After performing data quality checks, the researchers determined that while the population estimates were reasonable, local inaccuracies were still possible. For example, the data may not represent the complete evacuation of a specific neighbourhood or village.

**Geographic disaggregation and coverage:** data is available at the national, oblast, hromada, and 2km<sup>2</sup> cell levels.

**Age disaggregation:** five-year age groups until 80+, with the same limitations as the original University of Oxford dataset

**Sex disaggregation:** yes

**Update frequency:** daily for national and oblast levels and quarterly for hromada and 2km<sup>2</sup> cell levels

## World Bank

**Total population estimate as at 2022:** 38,000,000 (WB accessed 12/07/2023).

**Methodology:** the World Bank's estimate uses yearly extrapolations of new population figures from the SSSU, the UN's World Population Prospects and Population and Vital Statistics reports, and the Council of Europe (WB accessed 12/07/2023).

**Strength:** it includes additional data on rural/urban disaggregation and unemployment rates.

**Limitations:** it is unclear how the World Bank's different sources were combined to formulate the total population estimate, making it difficult to assess the data's limitations. Given the population estimate of 38,000,000 it is likely to be a post-invasion figure. The exact period in 2022 it refers to is unclear.

**Geographic disaggregation and coverage:** data is only available at the national level. There is no information on whether the data includes pre-February 2022 NGCAs, Crimea, and Sevastopol.

**Age disaggregation:** two age groups, 0–14 and 15–64

**Sex disaggregation:** yes

**Update frequency:** yearly

## IDPS AND RETURNEES

### IOM – general population survey

**IDP estimate as at June 2023:** 5,088,000 (IOM 03/07/2023)

**Estimated returns as at June 2023:** 4,757,000 (IOM 03/07/2023)

**Methodology:** from their initial survey of 20,000 respondents, IOM conducted follow-up phone interviews with 1,611 IDPs, 1,671 returnees, and 2,015 residents. The follow-up phone interview sample comprises the same population group distribution and oblasts as the initial survey (IOM 03/07/2023).

**Strengths:** the recently updated methodology uses a larger sample size. This dataset is the response's most commonly used source for IDP figures.

**Limitations:** the survey can only take place in areas where Ukrainian operators can provide mobile phone coverage, excluding territories under Russian control and potentially excluding areas heavily affected by the conflict. The proportion of the excluded population (people without access to mobile phones or not covered by Ukrainian mobile operators) is unknown. As a result, IDP estimates for Donetsk, Kherson, Luhansk, and Zaporizhzhia oblasts are likely to be lower than reality. The methodology also assumes that parents or guardians accompany minors (IOM 03/07/2023).

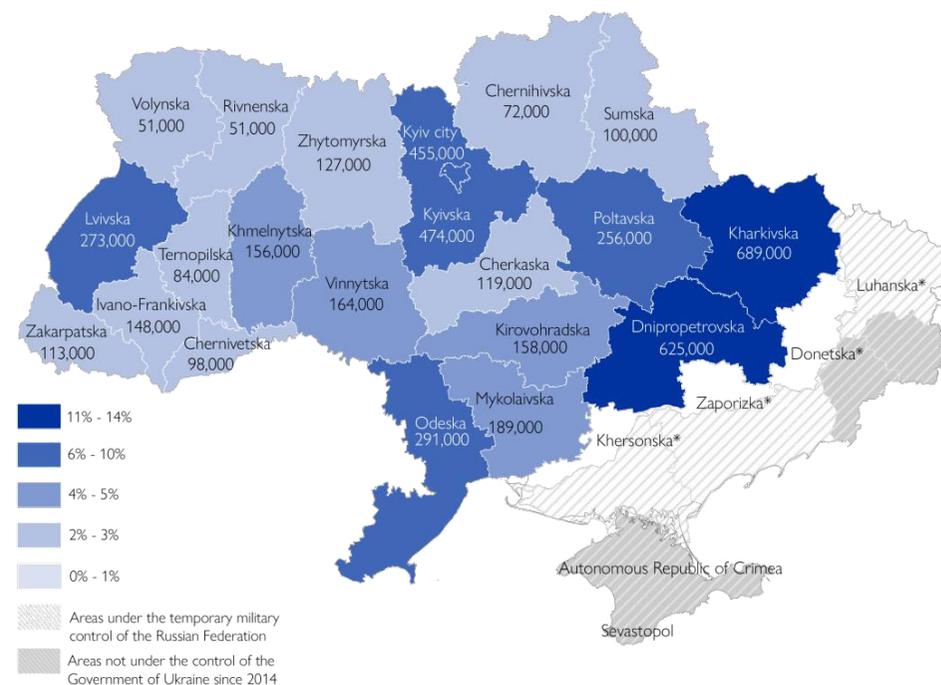
**Geographic disaggregation and coverage:** national and oblast data is available. Estimates do not cover Russian-controlled territories.

**Age disaggregation:** infants (0–1 year old); children (1–5, 6–10, and 11–17 years old); adults (18–29, 30–39, 40–49, and 50–59 years old), and older people (60+ years old)

**Sex disaggregation:** yes

**Update frequency:** quarterly from Round 13 onwards

### Estimated location of internally displaced persons by oblast



Source: IOM (03/07/2023)

## IOM – area baseline assessment

**Registered IDPs as at 30 June 2023:** 3,605,338 (IOM 14/07/2023)

**Methodology:** IOM collects recorded IDP figures from regional authorities via a network of key informants within oblast administration bodies (IOM 14/07/2023)

**Strengths:** it is the only data on IDPs available at the sub-oblast level and reflects official registration figures.

**Limitations:** the methodology relies on IDP registration. Not all IDPs have registered their displacement with Ukrainian authorities.

**Geographic disaggregation and coverage:** data is available down to the raion level. The data covers 78% of government-controlled hromadas (IOM 14/07/2023).

**Age disaggregation:** 0–18, 18–59, and 60+ years old

**Sex disaggregation:** yes

**Update frequency:** monthly

## University of Oxford

**Total publicly available IDP estimate as at July 2022:** 6,276,195 (Leasure et al. 11/05/2022)

**Methodology:** The data sums up the net population changes for oblasts whose July 2022 population estimate has decreased compared to the pre-invasion baseline. The pre-invasion baseline is taken from UNFPA's COD-PS dataset. The figure from the net population changes becomes the IDP estimate. It is worth noting that IOM's figures for July 2022 indicated 6,654,000 IDPs in Ukraine. This number is 5% more than the University of Oxford estimated for the same period, the University's most recent publicly available figures (IOM 29/07/2022).

**Strengths:** this report provides daily spatial and time granular data, allowing for trends analysis.

**Limitations:** the data cannot capture displacement within an oblast or whether displacement into an oblast offsets displacement from that oblast (Leasure et al. 11/05/2022).

**Geographic disaggregation and coverage:** there is no geographic disaggregation. It only provides the number of IDPs at the national level.

**Age disaggregation:** five-year age groups until 80+

**Gender disaggregation:** yes

**Update frequency:** daily until 1 July 2022 (Leasure et al. 11/05/2022).

University of Oxford's map depicting proportional net change in oblast population as at June 2022



Source: Leasure et al. (11/05/2022)