

# YEMEN HUMAN DEVELOPMENT SURVEY

## METHODOLOGICAL OVERVIEW



### EXECUTIVE SUMMARY

- › The Yemen Human Development Survey is the first face-to-face based dataset consisting of key development indicators representative of the Internationally Recognized Government (IRG) areas of Yemen since the start of the war.
- › The YHDS visited a sample of 1,681 households, 16 in each of 105 Enumeration Areas (EAs), themselves selected out of the 1,200 EAs visited by the National Yemen Household Budget Survey (HBS) in 2014. The sample is stratified by region, urban/rural location, and displacement status.
- › The YHDS collected data on seventeen distinct modules including: dwelling characteristics health, education, food security, displacement, coping strategies and access to social protection. Some modules were collected at the household level, while others collected detailed information on the individuals of the household.
- › The data quality control was a crucial element of the field work, given the limited ground presence and monitoring.
- › The response rate for the YHDS is 95%, which is relatively high considering the current security situation and extent of internal displacement in the country.
- › The YHDS indicates that, six percent of households have been displaced, and more than half have been exposed to high intensity conflict. The average household size is six, and a household is usually headed by a married man. There is a significant youth bulge, with two thirds of the IRG Yemeni population under the age of thirty.

**State and  
Peacebuilding  
Fund**



MANAGED BY  
**THE WORLD BANK**  
IBRD · IDA | WORLD BANK GROUP



## OBJECTIVE

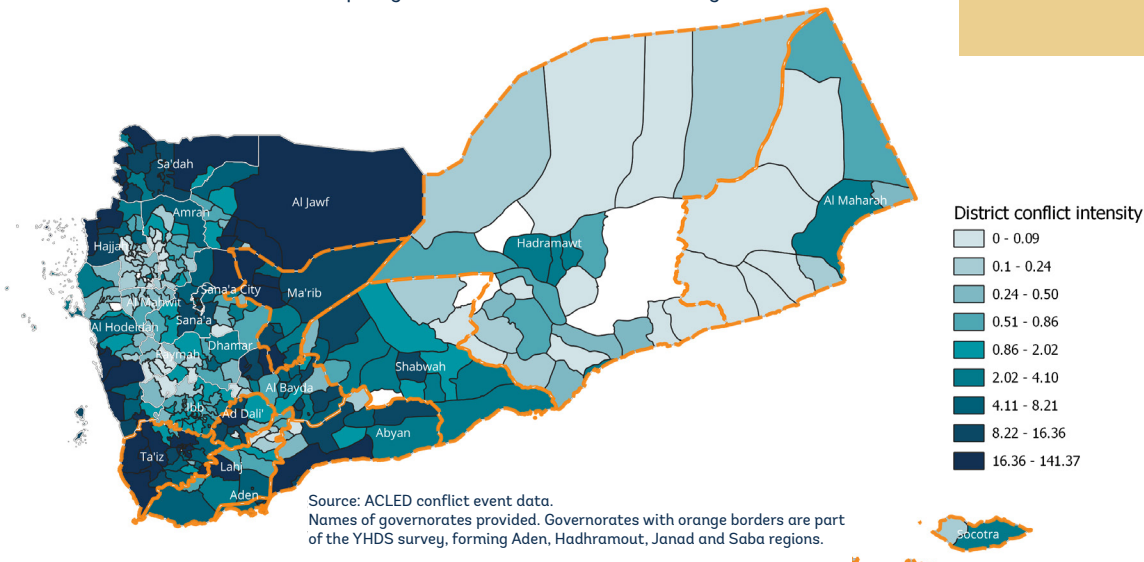
The Yemen Human Development Survey (YHDS) is the first face-to-face, representative household survey that seeks to fill several data gaps in understanding the state of human development indicators in IRG areas since the onset of conflict, as well as help understand the institutional landscape of human development service delivery. The survey was jointly designed by the World Bank and the Social Fund for Development, with inputs from UNDP. The YHDS seeks to provide a detailed overview of welfare, food security and human development indicators in Yemen, as well as critically enable intra-household analysis, to help identify needs and vulnerabilities by gender, across age groups, and between groups.

The survey was conducted from April to September of 2021, with a one-month break for the Ramadan period. The YHDS is a unique and comprehensive survey implemented by the Social Fund for Development in partnership with the World Bank, providing much needed information on the welfare of accessible areas of the Internationally Recognized Government (IRG) (or Southern Yemen).<sup>1</sup> The YHDS collects data on education, health, employment status and living conditions in IRG areas, as the country approaches seven years of a drawn out and devastating conflict. The YHDS offers regionally representative data drawn from the governorates of Abyan, Aden, Lahaj, Al-Daleh, Hadramout, Shabwah, Al Mahra, Taiz, and Marib.

The YHDS is the first dataset consisting of key development indicators representative of IRG areas of Yemen since the start of the war. While there have been several food security assessments over the last seven years, the YHDS offers a more comprehensive and richer picture of the living standards, economic wellbeing and human capital outcomes of Yemenis in IRG areas. This note explains the methodology used in designing, collecting and analyzing the YHDS. The resulting analysis is summarized in a series of thematic briefs that follow from this note.



YHDS sampled governorates and conflict intensity



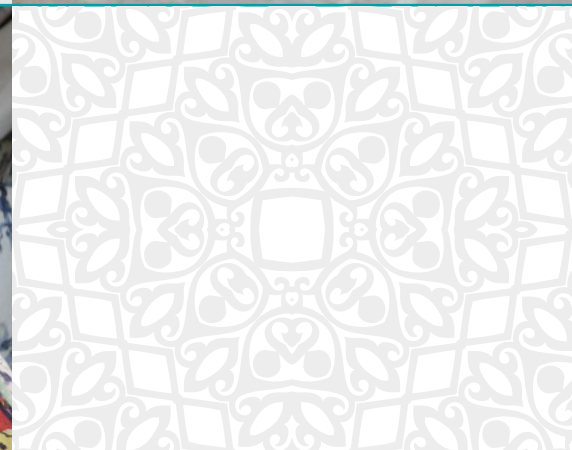
<sup>1</sup> The team was unable to gain access to the areas in the North of Yemen controlled by de-facto authorities (DFA).

## SAMPLE

The YHDS visited a sample of 1,681 households, 16 in each of 105 Enumeration Areas (EAs), themselves selected out of the 1,200 EAs visited by the National Yemen Household Budget Survey (HBS) in 2005 and 2014. The YHDS sample was effectively designed in four stages:

- › By design, the first stage is identical to that of the 2014 HBS, in which 1,200 EAs were selected from the list of all EAs generated by the 2005 Census frame using probability proportional to size.
- › In the second stage, a subset of 273 of the 1,200 HBS EAs were selected. The YHDS used the same 38 strata of the HBS (region and urban/rural), and considered as an additional, separate stratum, the districts where the ratio of incoming IDPs (as reported by the International Office for Migration (IOM)) over the total population (as reported by the Central Statistical Office (CSO)) exceeded 60%.
- › Due to difficulties obtaining security permissions in the Northern areas and increased hostilities in some districts, 168 EAs in the de facto areas were inaccessible. Out of the 105 EAs that are accessible in IRG areas, four needed to be replaced with reserve EAs once fieldwork began due to the ongoing security situation.
- › Finally, all households within the EAs were listed.<sup>2</sup> The households were sorted into those with IDPs and non-IDP households based on the following question in the household listing: “How many members of the current household have moved here because of the conflict?”. Then 8 IDP households and 8 non-IDP households were randomly selected with equal probability from each group.<sup>3</sup> In this way, an additional strata of IDP and non-IDP households is created.

Sampling weights were computed based on sampling probabilities which were adjusted using Stratum-wise factors to make the sum of weights match the number of households reported by the 2005 census within the scope of the survey. As such, statistics in reported from the YHDS are representative of the accessible parts of Southern Yemen or the accessible areas under the control of the IRG.



<sup>2</sup> There were five cases of large EAs, and the field team segmented the EA and randomly selected 2 segments. In these cases, households were listed from the 2 selected segments.

<sup>3</sup> If less than 8 households with IDPs are listed, a few more than 8 non-IDP households were selected, in order to always select 16 households in total.

Figure 1: Distribution of the EAs visited in YHDS 2021 and HBS 2014 by strata, with population estimates projected from 2005 census.

Region	Governorate	Estimated Population in 2017			Number of EAs in HBS 2014		No. of EAs visited in 2021		
		Districts with high IDP	Other districts		Urban	Rural	Districts with high IDP	Other districts	
			Urban	Rural				Urban	Rural
1 Aden	12 ابين Abyan		134,709	433,291	30	18	2	5	
	24 عدن Aden		925,000		72		10		
	25 لحج Lahaje		117,451	865,549	25	23	2	12	
	30 الضالع Al-Daleh		94,695	625,305	22	14	1	6	
2 Hadhramout	19 حضرموت Hadhramout	4,762	657,706	761,568	41	19	12	12	
	21 شبوه Shabwah		103,312	528,688	21	15	2	10	
	28 المهرة Al-Mahra		69,317	80,683	12	12	2	2	
3 Janad	11 إب Ibb		469,472	2,367,528	43	41			
	15 تعز Taiz		697,553	2,484,447	56	40	4	6	
4 Azal	13 الامانة Sana'a City		2,948,472		156				
	20 ذمار Dhamar		240,182	1,672,818	31	29			
	22 صعده Sad'dah		171,181	906,819	28	20			
	23 صنعاء Sana'a gov.		1,435,528		24				
	29 عمران Amran		177,436	874,564	27	21			
5 Saba	14 البيضاء Al-Bayda		148,058	611,942	29	19			
	16 الجوف Al-Jawf	36,126	84,190	468,684	22	14			
	26 مارب Marib	198,579	15,392	122,887	22	14	16	1	
6 Tahama	17 حجة Hajjah	254,612	162,606	1,711,782	30	30			
	18 الحديدة Al-Hudaydah		1,006,214	2,182,786	75	33			
	27 المحويت Al-Mahwit		46,319	648,681	27	21			
	31 ريمه Remah		566,000		24				
<b>Total</b>			<b>28,112,895</b>			<b>1,200</b>	<b>105</b>		

Note: Highlighted in yellow are the regions of the YHDS sample. In bold are the visited governorates.



## MODULES

The YHDS collected data on seventeen distinct modules. Some modules were collected at the household level, while others collected detailed information on individuals in the household.

When possible, the household questionnaire was administered to a female member of the household who is most knowledgeable about the household. As all the interviewers were female, this was also more acceptable to the respondents. In 72% of interviews the respondent was female, and almost half the respondents were the spouse of the head of the household. All respondents were older than fifteen to ensure that informed consent could be taken. The household level modules include dwelling characteristics, exposure to COVID-19, assets, food consumption and security, expenditure, income sources, problems and coping strategies, displacement, remittances, and social protection. Individual modules were responded to by the same respondent on behalf of other household members (response by proxy) unless they were available and able to answer the questions directly. The exception to this was the module on women in the household, which was necessarily answered by a female respondent. The individual modules include the household roster, health, education, employment, fertility, Under-5s, Children's time use, and women in the household modules.

Figure 2: YHDS modules

Module	Level	Detail	
1 	<b>Dwelling characteristics</b>	Household	Type of housing unit, number of rooms, rent, material used for walls, roof and floor of dwelling, access to and source of water, sanitation, and electricity access.
2 	<b>Exposure to Covid</b>	Household	Knowledge of Covid-19 symptoms, knowledge of precautions and precautions taken.
3 	<b>Household roster</b>	Individual	Sex, age, relationship to head, marital status, details of marriage, form of identification.
4 	<b>Health</b>	Individual	Disabilities, chronic diseases, acute illnesses and accidents, access to and status of health facilities, reasons for not seeking medical attention.
5 	<b>Education</b>	Individual	Literacy, current enrollment, highest level completed, reason for dropping out (for those not attending school and younger than 13), reason for never attending school, informal schooling, schooling expenses.
6 	<b>Employment</b>	Individual	Employment situation before and after the war, including: occupation, economic activity, wages, formality. For current status this module also asks about job search and underemployment.
7 	<b>Fertility</b>	Individual	Administered to ever married women in the household between the ages of 15 and 49. Includes information on number of births, infant deaths, pregnancy, access to antenatal care and birth attendance.
8 	<b>Under - 5s</b>	Individual	The main respondent is asked questions about all the children under 5 years in the household. Includes information on vaccinations, breastfeeding, malnutrition and diarrhea.
9 	<b>Children's time use</b>	Individual	Up to two children between the ages of 5 and 17 are randomly selected, and the main respondent is asked questions on the types of activities the child has done in the last week, including paid or unpaid work.
10 	<b>Household assets</b>	Household	Respondent is asked whether the household owned a set of 47 assets in addition to livestock, before 2015 and currently. The respondent is also asked why they saw a decrease in asset ownership and whether female household members own land and jewelry.
11 	<b>Food consumption</b>	Household	Household consumption of 17 food items over the last week as well as questions used to calculate the reduced form coping strategy score.
12 	<b>Expenditure</b>	Household	Household expenditure of 25 non-food items such as rent, electricity, transportation and entertainment.
13 	<b>Other income sources</b>	Household	Income received by any household member from sales or rent that are not included in the employment module, in addition to details on outstanding loans and credits.
14 	<b>Problems &amp; Coping strategies</b>	Household	Exposure of the household to a series of shocks, including natural disasters, theft, illness or violence, and the coping strategies used to recover from these shocks.
15 	<b>Displacement</b>	Household	Experience of migration over the course of the conflict, including the reasons for migration, the governorate/district of origin, experience hosting displaced households or IDPs, and intention to migrate.
16 	<b>Women in the household</b>	Individual	Administered to the female member who is either the wife of the head of household, the most active and important female member of the household, or in case of female headed household, the head of household. Includes information on household decision making, mobility to certain areas and the chaperones required, and feelings of safety.
17 	<b>Remittances</b>	Household	The amount and frequency of remittances sent and received from within and outside Yemen, as well as information on the relationship with the remitter or receiver.
18 	<b>Social protection</b>	Household	Household access to cash support, food support or any other type of support from a series of programs delivered by the UN, NGOs or CSOs.

## DATA COLLECTION

Data collection started on the 5th of April 2021 and ended on the 22nd of September 2021. Data was collected using Computer Assisted Personal Interviewing (CAPI), which aided in the efficiency of data collection and ensured that data could be stored regularly in a central database. The fieldwork was paused after one week over the Ramadan period, which provided a natural point to evaluate the quality of the data and debrief with the interviewers to clear up any concerns or questions. The field team consisted of forty female interviewers grouped into fourteen teams of approximately three interviewers led by a field supervisor. The teams conducted the household listing over two days before the households were randomly selected using a CAPI application and then interviewed over two days.



## DATA QUALITY CONTROL

The data quality control was a crucial element of the field work, given the limited ground presence and monitoring. For this reason, quality control was enforced in three ways:

1. Ensuring that the questionnaire was programmed using the CAPI software [ODK](#) such that any open-ended questions included reasonable limits (for example it should not be possible to work for more than 31 days per month) and inconsistencies in the answers were flagged to the interviewers (for example an individual younger than five cannot be the father of the household head). This level of quality control meant that interviewers were able to correct mistakes during the interview and clarify with the respondent if needed.
2. The second level consisted of monitoring data quality using an excel based dashboard. Anonymized (or deidentified) data was shared with the wider team on a regular basis (daily if possible, but sometimes every couple of days in case of internet issues). Key Performance Indicators (KPIs) for each interviewer were then produced in excel by running a set of Stata do-files. An example of these indicators include the average number of times the response recorded by the interviewer would trigger a set of subsequent questions (such as the household size), the idea being that an interviewer could misreport these responses with the intention of saving time. As seen in Figure 3 (names are removed from the excerpt for privacy reasons), interviewers who consistently had low filter responses or outliers in terms of the duration of the interview were flagged, and the field supervisor was asked to sit in on further interviews and take further action if needed. Outliers could also be identified from the dashboard, and these were clarified with the field team on a regular basis.
3. Finally, the field supervisors would attend interviews on an ad hoc basis, to ensure that the protocol was being followed and the interviewers were following the guidelines and instructions. In some cases, field supervisors were joined by data monitors from the SFD office.

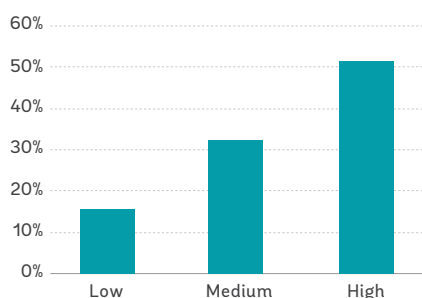
**Figure 3:** Excerpts from excel based dashboard used to monitor data quality



## RESPONSE RATE

Eighty-eight households did not complete the interview, either because they refused to take part or because they were unavailable despite several attempts to schedule an interview. This implies the response rate for the YHDS is 95%, which is relatively high considering the ongoing security situation in the country. Households who did not complete the interview were replaced by others to ensure a sufficient sample size was reached.

**Figure 4:** Percentage of households living in low, medium and high conflict intensity districts



Note: A conflict intensity score is calculated for each district using data from the Armed Conflict Location & Event Database. The district level score is a weighted indicator of the number of battles, explosions and conflict events in each district from 2015 until June 2021. A higher weight is given to events in more recent years. Categories of low, medium and high are determined such that the 333 districts of Yemen are divided into three equal groups in increasing order of their score. The district level data is then merged with the YHDS 2021.

## DESCRIPTION OF RESPONDENT HOUSEHOLDS

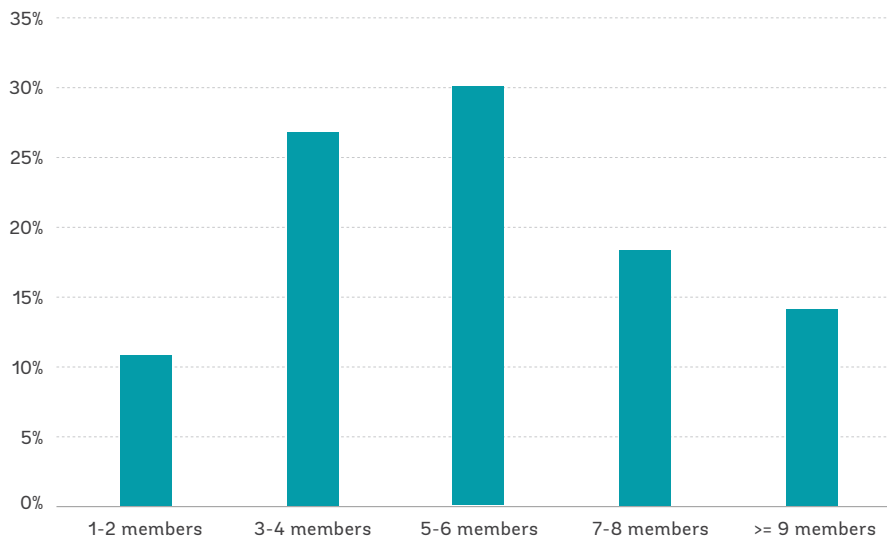
IRG areas of Yemen are predominantly rural, with 61 percent of households living in rural areas. Eight percent of households are currently displaced from their original homes because of the conflict and 17 percent have returned after being displaced due to conflict, implying 25 percent of households have ever been displaced.<sup>4</sup> Slightly more than half of the sample live in districts considered to be high intensity conflict, indicating the extent of conflict exposure experienced by most households in IRG areas.

The average number of household members in IRG areas is six. The average number of children per household (below 18) is 2.5 and 18 percent of households have at least one person over the age of 65. 17 percent of households are headed by a woman, and 86 percent of household heads are married. In cases where the household head is a woman, it is more likely that they are either divorced or widowed (56 percent of female headed HHs). 16 percent of the Yemeni population in IRG areas report having a disability.

<sup>4</sup> A roughly similar percentage of the population are currently displaced or have returned. Further details on the profiles of the displaced and returnees are examined in the policy brief on displacement.



Figure 5: Household size.



Source: YHDS 2021



When considering the individual level data, 51 percent of Yemenis in IRG areas are female and 37 percent are married. Almost half of the Yemeni population in IRG areas are younger than twenty, and two thirds are less than thirty years old, implying a large youth bulge.

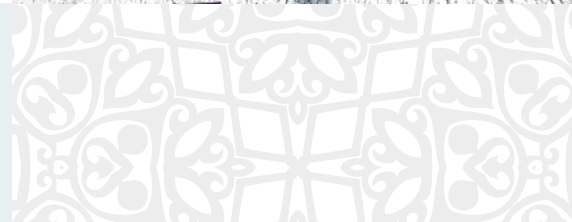
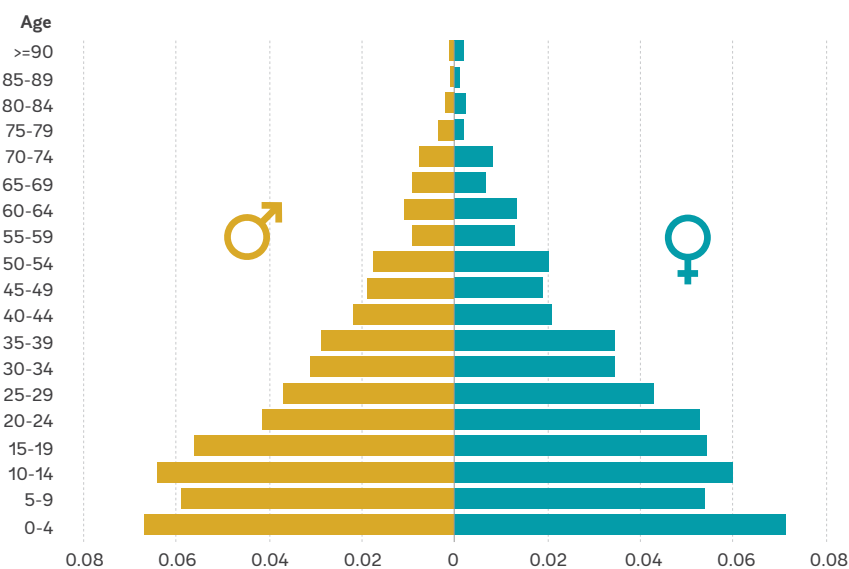


Figure 6: Population pyramid. Proportion of population by age group and gender.



Source: YHDS 2021



## CONCLUSION

The Yemen Human Development Survey offers valuable information on the lives and wellbeing of Yemeni households in IRG areas in a context of data scarcity and ongoing conflict. 1,681 households were interviewed over a period of six months, achieving a response rate of 95%. Through a series of thematic briefing notes accompanying this methodological overview, the YHDS provides much needed information on the living conditions and human development outcomes of Southern Yemenis living in the context of ongoing conflict and violence in IRG controlled areas. The accompanying briefing notes cover the sectors of education, health, women's empowerment, labor, social protection and remittances, and displacement.



