

Tajikistan 2023 Technical Document

Tajikistan Sampling

A stratified multi-stage cluster sample design was used to complete 1,001 face-to-face surveys.

Target Population/Coverage: Non-institutionalized adult population (15 years of age or older) living in households. Stratification and selection used information from the 2010 Population Census and Statistical Year Book 2018. The GBAO was excluded, as it was closed for any kind of surveys or field research by the national security service. The excluded region represents approximately 3% of the population.

Stratification: The sampling frame was stratified by geographic region and urban/rural status, resulting in a total of 7 strata groups. The regions used for stratification include: Sogdiyskaya, Khatlon, Dushanbe, and RRP. Dushanbe (the capital) is treated as a separate stratum, respectively. The regions were further stratified by urban and rural status.

Sample Selection: Primary Sampling Units (PSUs) are cities/towns, rural settlements (villages) and city districts (Dushanbe). PSUs were selected using probabilities proportional to population size, where the total (0+) population was the measure of size. A total of 100 ultimate clusters were selected in the final sample.

Within each selected PSU or SSU, a random starting point was identified. Using the right-hand rule and a fixed interval of 3 households (i.e. every third household is selected), interviewers selected eligible households until the required number of interviews were completed. In large apartment buildings (with 10 or more residential units), an interval of 5 was used. Within each selected household, interviewers listed all eligible (15+ adults) individuals and the CAPI program randomly selected a respondent. In case of refusals or non-contacts, households were replaced along the random route using the same interval.

Data Collection: October 7, 2023 – November 15, 2023

AAPOR3 response rate: 74%

Weighting: The sample data were weighted to minimize bias in survey-based estimates. The weighting procedure was formulated based on the sample design and was carried out in multiple stages. A probability weight factor (base weight) was constructed to correct for unequal selection probabilities. At the next step, the base weights were post-stratified to adjust for non-response and to match the weighted sample totals to known target population totals obtained from country-level census data.

Design Effect: 1.37

Margin of error: ± 3.6% (95% confidence level)

Population sources used for constructing weights were based on the following:

Age, Gender, Region, Urbanicity: Statistical Yearbook of the Republic of Tajikistan, 2018

Education: 2010 Population and Housing Census

Age	SAMPLE_UNWTD	SAMPLE_WTD	TARGET
15-19	9.8	14.0	14.1
20-29	21.9	28.2	28.3
30-39	21.8	22.4	22.5
40-49	17.6	14.6	14.6
50+	29.0	20.7	20.5
Education	SAMPLE_UNWTD	SAMPLE_WTD	TARGET
Incomplete sec or less	21.5	27.1	27.2
Secondary	46.0	54.8	55.0
Sec voca/incomplete higher/higher	32.6	18.1	17.8
Sex	SAMPLE_UNWTD	SAMPLE_WTD	TARGET
Male	41.8	49.8	50.0
Female	58.2	50.2	50.0
Region	SAMPLE_UNWTD	SAMPLE_WTD	TARGET
Sogdiyskaya oblast	31.0	30.4	30.4
Khatlon	35.1	35.4	35.4
Dushanbe	11.0	10.6	10.6
rrp	23.0	23.6	23.6
Urbanicity	SAMPLE_UNWTD	SAMPLE_WTD	TARGET
Urban	28.0	27.4	27.5
Rural	72.0	72.6	72.5