

01_test

Cal Chengqi Fang

2/19/2022

1. Sample Description

1.1 Sampling Strategy

The dataset we will use in this project is the outcome of the NCSS-SRH survey project. This project used a combination of multi-stage random sampling method and snowball convenience sampling method to collect the data.

In the first stage, researchers built the sampling frame based on the distribution of higher education schools in the eastern, central, and western regions, referring to the directory of higher education institutions published by the Ministry of Education of China in June 2018. Due to the severely uneven distribution of higher education institutions in China at the provincial (administrative region level equal to states in America) level (e.g., there are only seven colleges and universities in the whole Tibetan region), there are difficulties in sampling at the provincial level, and the stratification of the sample colleges and universities in the first sampling stage was performed using the East, Central and West divisions based on administrative divisions provided by the National Bureau of Statistics.

In the second stage, colleges offering associate's or bachelor's degree were selected from each region. Selected colleges offering bachelor's degrees were then divided into four levels in each region: first-class universities, universities of first-class disciplines, general undergraduate colleges and universities, and private undergraduate colleges and universities. Institutions offering associate's degrees were divided into three levels: key associate's colleges and universities, general associate's colleges and universities, and private associate's colleges and universities. Corresponding number of universities are selected according to the proportion of each level in the overall, with an appropriate inclination to the colleges and universities with more geographically diverse student source, taking into account the capability of the survey implementation organization, China Youth Network. A total of 241 sample colleges and universities are selected. The purpose of sampling at school level in this way is to ensure that the social networks in the study have nationwide first-level dissemination points, and to ensure that college students in 31 provinces and regions nationwide have access to the questionnaire through social media and can be included in the sample (in fact, college students in 34 provincial-level administrative regions, including Hong Kong, Macao and Taiwan, are included in the snowballing process). However, the samples from Hong Kong, Macao and Taiwan were suggested to be excluded in the analysis stage because of the significant differences in social background, cultural environment and education system between these samples and those from mainland China.

The 241 sample universities were used as the starting point for snowballing to sample college students enrolled in higher education institutions nationwide. All samples were *volunteer samples* recruited in the form of convenience sampling, with students deciding whether or not to participate in the survey and being able to withdraw at any time during the process after participation, with full disclosure of survey ethics given in the guidelines.

Referring to the literature review, we could see that most of the similar existing researches on youth sexual knowledge, attitudes, and behavior conducted in other Asian countries used random sample. However,

