**IMPORTANT:** This assignment is only for those participants working in national statistical office

**Assignment II: Subsetting census data**

This assignment is designed for statisticians working in national statistical offices who are engaged in production and/or dissemination of population census data. It is expected to help them better understand census dataset, variables required for conducting thematic analysis on youth and extracting relevant tables from census micro-data that can be used for further analysis. This is also a preparatory work for those participants who are selected for the face-to-face phase of this training.

IMPORTANT NOTICE: *Participants are requested to bring the sample that they have selected from their population census (output from step 1 under instructions) to the face-to-face phase of the course. The dataset is necessary for successful completion of the data workshop as part of the training course. The sample data set has to include all the census questionnaire items.*

***Instructions***:

1. Calculate sampling rate for your country as $R=\frac{15000000}{N}$ in which *N* is your population according to the latest census.
2. Select a ***R%*** random sample of “***households***” from complete list of household in your population census dataset within each major administrative division (e. g. Province, State, District, Region, etc) separately for rural and urban areas. For example, if your largest administrative region is province and your country has 20 provinces, select *R*% of urban households and *R*% of rural households in each region.

***Example:*** *Think about a country with population of 15,000,000 (and 3,000,000 households). According to step (1),* $R=\frac{15000000}{15000000}=1$***.*** *Assume that province A has 200,000 urban households and 100,000 rural households, you need to select 1% random sample from urban and rural households (0.01\*200,000=2000 urban households and 0.01\*100,000=1000 rural households).*

1. Extract information on all census items for “***all individuals within selected households***” and save the new dataset in a separate file with excel format and name it ***subset.xls***
2. Use ***subset.xls*** dataset and extract information requested in tables under sections I, II and III in the template (below).
3. Insert the information from step (4) into tables provided in the template and submit your assignment.

Those who will attend face-to-face phase of the training: “Don’t forget to bring ***subset.xls*** with you to the course!”

**Template for the report on assignment II for statisticians:**

**Participant’s name: ……………..**

1. **Table on fertility statistics**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Number of children ever born alive | Total female population |
|  | **Age of mother** | Male | Female | Total |
|  |  **10 - 14** | …….. | …….. | …….. | …….. |
| Youth |  **15 - 19** | …….. | …….. | …….. | …….. |
|  **20 - 24** | …….. | …….. | …….. | …….. |
| **Subtotal (15 - 24)** | …….. | …….. | …….. | …….. |
|  |  **25 - 50** | …….. | …….. | …….. | …….. |
|   | **Total** | …….. | …….. | …….. | …….. |

1. **Table on education statistics**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Population attending school | Total population |
|  | **Age** | Male | Female | Total | Male | Female |
| **5 - 9** | …….. | …….. | …….. | …….. | …….. |
|  |  **10 - 14** | …….. | …….. | …….. | …….. | …….. |
| Youth |  **15 - 19** | …….. | …….. | …….. | …….. | …….. |
|  **20 - 24** | …….. | …….. | …….. | …….. | …….. |
| **Subtotal (15 - 24)** | …….. | …….. | …….. | …….. | …….. |
|  |  **25 - 29** | …….. | …….. | …….. | …….. | …….. |
|   | **Total** | …….. | …….. | …….. | …….. | …….. |

1. **Table on employment statistics**

**`**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Active population | Total population |
|  |  | Male | Female | Male | Female |
|  | **Age** | Unemployed | Employed | Unemployed | Employed |
|  |  **10 - 14** | …….. | …….. | …….. | …….. | …….. | …….. |
| Youth |  **15 - 19** | …….. | …….. | …….. | …….. | …….. | …….. |
|  **20 - 24** | …….. | …….. | …….. | …….. | …….. | …….. |
| **Subtotal (15 - 24)** | …….. | …….. | …….. | …….. | …….. | …….. |
|  |  **25 +** | …….. | …….. | …….. | …….. | …….. | …….. |
|   | **Total** | …….. | …….. | …….. | …….. | …….. | …….. |