**8. A Project in Designing an Area Sample**

In this project you will design a sample of census tracts, block groups, and persons from Anne Arundel county in the state of Maryland in the U.S. Based on considering analytic subgroups, the desired precision of estimates, and the available budget, it has been determined that these sample sizes are to be selected:

|  |  |
| --- | --- |
| Age group | Sample sizes |
| 18-24 years | 200 |
| 25-44 years | 200 |
| 45-54 years | 200 |
| 55-64 years | 200 |
| 65+ | 200 |
| Total | 1000 |
|  |  |
| Sample tracts | 25 |
| Sample block groups per tract | 1 |

The sample design will use tracts as PSUs, block groups as SSUs, and persons as elements. The goals of the sample design are to select a sample of the sizes above while (i) achieving a self-weighting sample in each of the age groups above and (ii) obtaining an equal workload in each sample PSU. You should pay particular attention to geographic areas that have small population counts and decide how they should be handled in the frame.

Use Sampford’s method to select the PSUs and SSUs. This is one of several options for selecting probability proportional to size samples. Sampford works for samples of any size and permits joint selection probabilities to be computed. This method of selection is available in R pps and sampling packages and in SAS proc surveyselect. In order to reproduce the solution given later in Chapter 12, include the statement

set.seed(-741881304)

at the beginning of your program if you use R. If you use SAS surveyselect, use the option

seed = 1953.

The deliverables for the project will be

* A sampling report
* SAS or text files giving the units used for the area frame and relevant census counts and measures of size
* SAS or text file for the selected sample along with relevant census counts, measures of size, selection probabilities, and weights.

**Contents of the Sampling Report**

Below is a list of topic areas that should be included in your report. The order of the sections in your report does not have to be the same as that given below. You should construct your report in a way that presents topics in an order that seems logical to your team.

The report should be written to a client whose staff includes managers and technical personnel. Managers will be more interested in understanding the broad outline of the steps used in weighting. Technical personnel will be interested in understanding the details of sample selection and weight computation, including appropriate formulae. You should consider how to structure your report to serve these audiences.

Topic Areas for the Sampling Report

Title Page (project title, date of submission, and name of project contact person)

Introduction (overview of the document)

Sample Design

* Goals of the sample design
* Area sampling frame
* units, data available, source of the data
* Assigning measures of size to units

Sample Selection

* Method of selection
* Selected units and characteristics of each
* Selection probabilities of units at each stage of selection
* Description of how persons should be selected from area listings

Maps

* Anne Arundel county
* Selected tracts and block groups

Appendix

* PROC CONTENTS or codebook of frame and sample files
* Listing of the sample PSUs and sample SSUs with their selection probabilities and census data. On each sample SSU, list the sampling rate you will use to select persons in each domain.

**Data Files and Other Information**

*AnneArundel.MD.xls*—Census 2000 tract and block group data for Anne Arundel county

*Census.glossry2.pdf*—defines geographic terms used by Census Bureau

Census tract and block maps for Maryland from the Census Bureau

American Factfinder at Census Bureau: factfinder.census.gov

[www.census.gov/geo/www/maps/CP\_MapProducts.htm](http://www.census.gov/geo/www/maps/CP_MapProducts.htm)

Maps of the county are also in

*Anne Arundel.blkgrps(streets).pdf*

*Anne Arundel.tracts(streets).pdf*

*Anne Arundel.tracts(no streets).pdf*