**f**

1. **Project 1: Design a Single-stage Personnel Survey**

Our primary goal is to equip survey researchers with the tools needed to design and weight survey samples. This chapter gives the first of several projects that mirror some of the complexities found in applied work. The three goals of this project are:

* Determine the allocation of a single-stage sample to strata in a multipurpose survey, accounting for specified precision targets for different estimates and differing eligibility and response rates for subgroups;
* Examine how sensitive the precision of estimates is to incorrect assumptions about response rates; and
* Write a technical report describing the sample design.

As you proceed through the following chapters in Part I of the book, we suggest that you return to this chapter periodically, refresh your memory about the aims of Project 1, and think about how the methods in Chapters 3-6 can be used in the development of the sampling design.

* 1. **Specification for the Study**

The Verkeer NetUltraValid (VNUV) International Corporation is preparing to conduct Cycle 5 of its yearly climate survey of employees in their Survey Division. The climate survey assesses employee satisfaction in various areas such as day-to-day work life, performance evaluations, and benefits. In the first three cycles of the survey, the VNUV Senior Council attempted to do a census of all employees, but many employees considered the survey to be burdensome and a nuisance. The response rates progressively declined over the first three cycles. In the fourth cycle, the Senior Council decided to administer an intranet survey only to a random sample of employees within the Survey Division The aim was to control the sampling so that continuing employees would not be asked to respond to every survey. In Cycle 5, a more efficient sample is desired that will improve estimates for certain groups of employees. The Senior Council requires a report from your design team that specifies the total number of employees to be selected, as well as their distribution by a set of characteristics noted below. They wish the quality and precision of the estimates to be better than the Cycle 4 survey. Note that this is the first survey in which the Senior Council has sought direction from sampling statisticians on the allocation of the sample.

Three business units are contained in the Survey Division: (*i*) the Survey Research Unit (SR) houses both survey statisticians and survey methodologists; (*ii*) the Computing Research Unit (CR) contains programmers who support analytic and data collection tasks; and (*iii*) Field Operations (FO) is populated by data collection specialists. The Senior Council would like to assess the climate within and across the units, as well as estimates by the three major salary grades (A1-A3, R1-R5, and M1-M3) and by tenure (i.e., number of months employed) within the units. However, the climate survey will only be administered to full- and part-time employees within these units. Temporary employees and contractors are excluded from the survey.

The Senior Council has identified three questions from the survey instrument that are most important to assessing the employee climate at VNUV. They are interested in the percentages of employees answering either “strongly agree” or “agree” to the following questions:

Q5. Overall, I am satisfied with VNUV as an employer at the present time.

Q12. There is a clear link between my job performance and my pay at VNUV.

Q15. Overall, I think I am paid fairly compared with people in other organizations who hold jobs similar to mine.

Note that the response options will remain the same as in previous years. Namely, a five-level Likert scale: Strongly Agree, Agree, Neutral, Disagree, and Strongly Disagree. A sixth response option, Don’t know/Not Applicable, is also available.

Additionally, the Senior Council would like estimates of the average number of training classes attended by the employees in the past 12 months. Relevant classes include lunch-time presentations, formal instructional classes taught at VNUV, and semester-long courses taught at the local universities.

* 1. **Questions Posed by the Design Team**

After receiving the study specifications document from the Senior Council, a design team is convened to discuss the steps required to complete the assigned task. At this initial meeting, the following information was determined from the specifications.

* Data will be collected from employees through a self-administered intranet (i.e., Web site internal to the corporation) questionnaire.
* All full- and part-time employees in the three business units within the Survey Division are eligible for the survey. Employees in other units within VNUV, as well as temporary employees and contractors, are ineligible and will be excluded from the sampling frame.
* The sample of participants will be randomly selected from a personnel list of all study-eligible employees provided by the head of VNUV’s Human Resources (HR) Department.
* A single-stage stratified sampling design is proposed for the survey because (*i*) study participants can be selected directly from the complete HR (list) sampling frame, and (*ii*) estimates are required for certain groups of employees within VNUV.
* The stratifying variables will include *business unit* (SR, CR, and FO), *salary grade* (A1-A3, R1-R5, and M1-M3), and potentially a categorized version of *tenure*.
* The analysis variables used to determine the allocation include three proportions, corresponding to each of the identified survey questions, and one quantitative variable. Estimates from the previous climate survey will be calculated by the design team from the analysis data file maintained by HR.

As is often the case when reviewing a sponsor’s specifications for a project, there were a number of issues that needed clarification. Based on the initial discussion, the design team submitted the following clarifying questions to the Senior Council and received the responses noted below each.

1. Currently, HR defines tenure as the number of months of employment at VNUV. Is there a grouping of tenure years that would be informative to the analysis? For example, analysis of the previous climate survey suggests that responses differ among employees with less than 5 years of employment at VNUV in comparison to those with a longer tenure.

Response: *Yes. Dichotomize tenure by less than 5 years and 5 years or greater.*

1. What is the budget for the climate survey and should we consider the budget when deciding on the total sample size?

Response: *The budget permits two staff members to be assigned part-time to process and analyze the data over a three-month period. This does not affect the sample size. However, the council has decided that individual employees should not be surveyed every cycle to reduce burden and attempt to get better cooperation. Selecting a sample large enough to obtain 600 respondents will permit the annual samples to be rotated among employees.*

1. We are interested in classifying a difference between two estimates as being substantively meaningful to VNUV. Could you provide us with a meaningful difference?

Response: *At least a five percentage point difference between any two sets of employee climate estimates is a meaningful difference. A 2-3 day difference in the average number of training classes is also of interest.*

1. Should the proportion answering “strongly agree” or “agree” to the three questions include or exclude the “don’t know/not applicable” response category?

Response: *Exclude.*

1. How precise should individual estimates be for this round of the survey? The quality of the data from prior versions of the climate survey has been measured in terms of estimated coefficients of variation (CV).

Response: *The target CVs of overall estimates by business unit, by tenure within business unit, and by salary grade within business unit are listed in Table 2.6 below.*

1. Are there additional requirements for the design, such as estimates by gender, by number of dependents, etc. in addition to estimates by business unit, business unit by salary grade, and business unit by tenure?

Response: *No.*

1. The VNUV Climate Survey Cycle 4 report does not detail the previous sampling design. The design team assumes that the Cycle 4 sample was randomly drawn from an updated list of employees within certain employee subgroups (i.e., a stratified simple random sample design). Is this correct? If so, where might we locate the stratifying information?

Response: *No strata were used in the last design. The previous employee file was sorted by a random number and an equal probability sample was selected.*

1. Are the eligibility and response rates expected to be the same in Cycle 5 as they were in Cycle 4?

Response: *The eligibility rates should be about the same, but we are not sure about the response rates. We would like to understand how sensitive the CVs will be if the response rates turn out to be lower than the ones in Cycle 4.*

* 1. **Preliminary Analysis**

HR provided the team with two data files. The first file contained information on all current VNUV employees such as employee ID, division, business unit, tenure in months, part-time/full-time status, and temporary/permanent employee status. The team eliminated all records for employees currently known to be ineligible for the survey, created a dichotomized version of tenure, and calculated population counts for the 18 design strata (Table 2.1).

The second file contained one record for each employee selected for the previous climate survey. In addition to the survey status codes (ineligible, eligible respondent, and eligible nonrespondent) and the survey responses, this file included the characteristics that should be used to define sampling strata in the new survey. This file, however, did not contain employee names or other identifying information to maintain the confidentiality promised to all survey participants. Sample members were classified as ineligible if, for example, they had transferred to another business unit within VNUV or retired after the sample was selected but before the survey was administered. The team isolated the eligible Survey Division records, created the sampling strata defined for the current climate survey design, and created the binary analysis variables for Q5, Q12, and Q15 from the original five-category questions (Table 2.2).

**Table 2.1 Distribution of Eligible Employees by Business Unit, Salary Grade, and Tenure:**

**VNUV Climate Survey Cycle 5, Survey Division**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **Business Unit** | | |  |
| **Salary Grade** | **Tenure** | **SR** | **CR** | **FO** | **Total** |
| A1-A3 | Less than 5 Years | 30 | 118 | 230 | 378 |
|  | 5+ Years | 44 | 89 | 115 | 248 |
| R1-R5 | Less than 5 Years | 106 | 86 | 322 | 514 |
|  | 5+ Years | 253 | 73 | 136 | 462 |
| M1-M3 | Less than 5 Years | 77 | 12 | 48 | 137 |
|  | 5+ Years | 44 | 40 | 46 | 130 |
| A1-A3 | *Total* | 74 | 207 | 345 | 626 |
| R1-R5 | *Total* | 359 | 159 | 458 | 976 |
| M1-M3 | *Total* | 121 | 52 | 94 | 267 |
| *Total* | Less than 5 Years | 213 | 216 | 600 | 1,029 |
|  | 5+ Years | 341 | 202 | 297 | 840 |
| *Total* | *Total* | 554 | 418 | 897 | 1,869 |

**Table 2.2 Documentation for Recode of Question Responses to Binary Analysis Variable:**

**VNUV Climate Survey Cycle 4, Survey Division**

|  |  |
| --- | --- |
| **Question Responses** | **Binary Analysis Variable** |
| 1 = Strongly Agree | 1 = Strongly Agrees or Agrees |
| 2 = Agree | 1 = Strongly Agrees or Agrees |
| 3 = Neutral | 0 = Does not (Strongly) Agree |
| 4 = Disagree | 0 = Does not (Strongly) Agree |
| 5 = Strongly Disagree | 0 = Does not (Strongly) Agree |
| 6 = Don’t know/Not Applicable | < missing category > |

The information in Tables 2.3-2.6 was tabulated from the Survey Division responses to the Cycle 4 survey. No survey weights were used because the Cycle 4 sample employees were selected with equal probability and no weight adjustments, e.g., for nonresponse, were made.

**Table 2.3 Distribution of Response Status by Business Unit, Salary Grade, and Tenure:**

**VNUV Climate Survey Cycle 4, Survey Division**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **Total** |  |  | **Eligible** | | | | |
| **Business** | **Salary** |  | **Sample** | **Ineligible** *a* | | **Total** | **Respondent** | | **Nonrespondent** | |
| **Unit** | **Grade** | **Tenure** | **‘n** | **‘n** | **pct** *b* | **‘n** | **‘n** | **pct** *c* | **‘n** | **pct** *c* |
| SR | A1-A3 | Less than 5 Years | 10 | 0 | 0.0 | 10 | 9 | 88.9 | 1 | 11.1 |
|  |  | 5+ Years | 11 | 0 | 0.0 | 11 | 9 | 84.6 | 2 | 15.4 |
|  | R1-R5 | Less than 5 Years | 34 | 3 | 9.7 | 31 | 16 | 51.6 | 15 | 48.4 |
|  |  | 5+ Years | 71 | 1 | 1.3 | 70 | 55 | 78.7 | 15 | 21.3 |
|  | M1-M3 | Less than 5 Years | 23 | 0 | 0.0 | 23 | 21 | 91.3 | 2 | 8.7 |
|  |  | 5+ Years | 13 | 2 | 15.4 | 11 | 9 | 84.6 | 2 | 15.4 |
| CR | A1-A3 | Less than 5 Years | 41 | 3 | 7.1 | 38 | 22 | 58.6 | 16 | 41.4 |
|  |  | 5+ Years | 20 | 0 | 0.0 | 20 | 10 | 50.0 | 10 | 50.0 |
|  | R1-R5 | Less than 5 Years | 28 | 0 | 0.0 | 28 | 14 | 50.0 | 14 | 50.0 |
|  |  | 5+ Years | 19 | 0 | 0.0 | 19 | 10 | 53.8 | 9 | 46.2 |
|  | M1-M3 | Less than 5 Years | 6 | 0 | 0.0 | 6 | 6 | 100.0 | 0 | 0.0 |
|  |  | 5+ Years | 9 | 1 | 11.1 | 8 | 7 | 88.9 | 1 | 11.1 |
| FO | A1-A3 | Less than 5 Years | 85 | 26 | 30.3 | 59 | 23 | 39.4 | 36 | 60.6 |
|  |  | 5+ Years | 16 | 0 | 0.0 | 16 | 6 | 39.4 | 10 | 60.6 |
|  | R1-R5 | Less than 5 Years | 101 | 2 | 2.2 | 99 | 65 | 65.2 | 34 | 34.8 |
|  |  | 5+ Years | 34 | 1 | 2.6 | 33 | 24 | 71.8 | 9 | 28.2 |
|  | M1-M3 | Less than 5 Years | 14 | 0 | 0.0 | 14 | 14 | 100.0 | 0 | 0.0 |
|  |  | 5+ Years | 14 | 2 | 15.4 | 12 | 10 | 84.6 | 2 | 15.4 |
| Total |  |  | 549 | 41 | 7.5 | 508 | 330 | 65.0 | 178 | 35.0 |

*a* Ineligible sample members were those employees selected for the Cycle 4 survey that retired or left the company prior to data collection.

*b* Unweighted percent of total sample within each design stratum (row).

*c* Unweighted percent of total eligible sample within each design stratum (row).

**Table 2.4 Estimates for Four Key Questions by Business Unit, Salary Grade, and Tenure:**

**VNUV Climate Survey Cycle 4, Survey Division**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Business**  **Unit** | **Salary**  **Grade** |  | **Proportion (Strongly) Agree** | | | **Avg Number of Training Classes** | |
| **Tenure** | **Q5** | **Q12** | **Q15** | **Mean** | **SE***a* |
| SR | A1-A3 | Less than 5 Years | 0.93 | 0.88 | 0.77 | 8.2 | 0.82 |
|  |  | 5+ Years | 0.75 | 0.71 | 0.62 | 12.4 | 1.24 |
|  | R1-R5 | Less than 5 Years | 0.84 | 0.80 | 0.69 | 22.3 | 2.23 |
|  |  | 5+ Years | 0.80 | 0.76 | 0.66 | 24.0 | 1.92 |
|  | M1-M3 | Less than 5 Years | 0.91 | 0.86 | 0.75 | 8.3 | 0.83 |
|  |  | 5+ Years | 0.95 | 0.90 | 0.79 | 3.6 | 0.36 |
| CR | A1-A3 | Less than 5 Years | 0.99 | 0.94 | 0.92 | 7.2 | 0.87 |
|  |  | 5+ Years | 0.80 | 0.76 | 0.74 | 10.9 | 1.09 |
|  | R1-R5 | Less than 5 Years | 0.82 | 0.78 | 0.76 | 19.6 | 3.92 |
|  |  | 5+ Years | 0.90 | 0.86 | 0.84 | 21.1 | 2.11 |
|  | M1-M3 | Less than 5 Years | 0.97 | 0.92 | 0.90 | 7.3 | 0.73 |
|  |  | 5+ Years | 0.97 | 0.92 | 0.90 | 3.2 | 0.32 |
| FO | A1-A3 | Less than 5 Years | 0.50 | 0.48 | 0.45 | 4.6 | 0.69 |
|  |  | 5+ Years | 0.52 | 0.49 | 0.47 | 6.9 | 1.04 |
|  | R1-R5 | Less than 5 Years | 0.75 | 0.71 | 0.68 | 12.5 | 1.87 |
|  |  | 5+ Years | 0.70 | 0.67 | 0.63 | 13.4 | 2.02 |
|  | M1-M3 | Less than 5 Years | 0.93 | 0.88 | 0.84 | 4.6 | 0.70 |
|  |  | 5+ Years | 0.94 | 0.89 | 0.85 | 2.0 | 0.30 |
| *a Standard error.* | | | | | | | |

**Table 2.5. Estimates By Business Unit, Salary Grade, and Tenure:VNUV Climate Survey Cycle 4, Survey Division**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Business**  **Unit** | **Salary**  **Grade** | **Tenure** | **Proportion (Strongly) Agree** | | | **Avg Number of Training Classes** | |
| **Q5** | **Q12** | **Q15** | **Mean** | **SE** |
| SR |  |  | 0.84 | 0.80 | 0.69 | 18.1 | 0.98 |
| CR |  |  | 0.90 | 0.85 | 0.83 | 12.6 | 0.90 |
| FO |  |  | 0.67 | 0.63 | 0.60 | 8.9 | 0.60 |
| SR | A1-A3 |  | 0.82 | 0.78 | 0.68 | 10.7 | 0.65 |
|  | R1-R5 |  | 0.81 | 0.77 | 0.67 | 23.5 | 2.26 |
|  | M1-M3 |  | 0.92 | 0.88 | 0.76 | 6.6 | 0.30 |
| CR | A1-A3 |  | 0.91 | 0.86 | 0.85 | 8.8 | 0.46 |
|  | R1-R5 |  | 0.86 | 0.81 | 0.80 | 20.3 | 5.45 |
|  | M1-M3 |  | 0.97 | 0.92 | 0.90 | 4.1 | 0.09 |
| FO | A1-A3 |  | 0.51 | 0.48 | 0.46 | 5.4 | 0.33 |
|  | R1-R5 |  | 0.74 | 0.70 | 0.66 | 12.8 | 2.09 |
|  | M1-M3 |  | 0.93 | 0.89 | 0.84 | 3.4 | 0.15 |
| SR |  | Less than 5 Years | 0.88 | 0.83 | 0.73 | 15.3 | 1.33 |
|  |  | 5+ Years | 0.81 | 0.77 | 0.67 | 19.9 | 2.06 |
| CR |  | Less than 5 Years | 0.92 | 0.88 | 0.86 | 12.2 | 2.67 |
|  |  | 5+ Years | 0.87 | 0.83 | 0.81 | 13.1 | 0.82 |
| FO |  | Less than 5 Years | 0.67 | 0.64 | 0.60 | 8.8 | 1.08 |
|  |  | 5+ Years | 0.67 | 0.63 | 0.60 | 9.2 | 1.02 |

**Table 2.6. Target Coefficient of Variation by Reporting Domain:**

**VNUV Climate Survey Cycle 5, Survey Division.**

|  |  |
| --- | --- |
| **Reporting Domain** | **Target CV***a* |
| Business Unit | 0.06 |
| Unit × Salary Grade | 0.10 |
| Unit × Tenure | 0.10 |

*a* Coefficient of variation.

* 1. **Documentation**

With the preliminary analysis complete, the design team began to draft the sampling report to the Senior Council using the following annotated outline:

Title = *UNUV Climate Survey Cycle 5 Sample Design Report*

1. Executive Summary
   1. Provide a brief overview of the survey including information related to general study goals and year when annual survey was first implemented
   2. Describe the purpose of this Cycle 5 document
   3. Provide a table of the sample size to be selected per business unit (i.e., respondent sample size inflated for ineligibility and nonresponse)
   4. Overview of the contents of the remaining section of the report.
2. Sample Design

* Describe the target population for Cycle 5
* Describe the sampling frame including the date and source database
* Describe the type of sample and method of sample selection to be useed

1. Sample Size and Allocation
   1. Optimization Requirements

* Optimization details including constraints, budget, etc.
* Detail the minimum domain sizes and mechanics used to determine the sizes
  1. Optimization Results
* Results = minimum respondent sample size per stratum
* Marginal sample sizes for key reporting domains
* Estimated precision achieved by optimization results
  1. Inflation Adjustments to allocation solution
* Nonresponse adjustments
* Adjustments for ineligible sample members
  1. Final Sample Allocation
* Marginal sample sizes for key reporting domains
  1. Sensitivity Analysis
* Results from comparing deviations to allocation after introducing changes to the optimization system

1. Appendix

* Sample size per strata (table), full sample and expected number of respondents
* Other relevant detailed tables including preliminary analysis
  1. **Next Steps**

The optimization problem and a proposed solution to the sampling design task discussed in this chapter will be revealed in Chapter 7. The methods discussed in the interim chapters will provide you with the tools to solve the allocation problem yourself. We will periodically revisit the VNUV design team discussions prior to Chapter 7 to provide insight into the design team’s decisions and procedures.