Customer Satisfaction Surveys

Implementation Manual

USAEP Water Indicators for Satisfaction Evaluation (*WISE*) Project
Research Triangle Institute
May 2000

Stephen Dunn (RTI, Team Leader)
Tarcicius Isbandhi (Consultant, Local Team Leader)
Eddy Akhirwan (Consultant/PERPAMSI, Water Sector Survey Specialist)
Winarko Hadi (Consultant/IATPI, Water Sector Survey Specialist)

Contents

Contents i

Acknowledgements ii

1. Overview of CSS Implementation Manual 1

2. Activities Required to Implement a CSS 3

3. Survey Instruments 6
3.1 Questionnaire Design 6
3.2 Question Design 10
3.4 Respondent Identification Number 13

4. Survey Field Operations 15
4.1 Pre Morning Meeting with Enumerators 16
4.2 Morning Meeting with Enumerators 16
4.3 Intra-day Monitoring and Support of Enumerators 17
4.4 Evening Meeting with Enumerators 18
4.5 Post Evening Meeting Activities 18

5. Enumerators 19
5.1 Enumerator Selection 19
5.2 Enumerator Training 21
5.2.1 Outline of Enumerator Training 21
5.3 Enumerator Monitoring 22
5.4 Enumerator Behavior and Compensation 23

6. Sampling 24
6.1 Brief Overview of Sampling Concepts 24
6.2 Sample Size 25
6.3 Basic Sampling Frame Concepts 29
6.4 Stratified Random Sampling 30
6.5 Sampling in the Absence of a Proper Sampling Frame 32

7. Data Entry and Cleaning 34
7.1 Data Entry 34
7.2 Data Cleaning 35

8. Data Analysis 37
8.1 Overview of CSS Analysis 37
8.2 An Illustration of CSS Analysis in Kotamadya Havana 37
8.2.1 Objectives of CSS Analysis 38
8.2.2 Illustrative Analysis Process and Results 39

9. Uses of Survey Data 46
9.1 Internal Uses of CSS Results 46
9.2 External Uses of CSS Results 48

Acknowledgements

The participation of many individuals and groups made possible the production of this *Customer Satisfaction Survey Implementation Manual*.

Specifically, the *Water Indicators for Satisfaction Evaluation* project team is grateful to:

- The United States-Asia Environmental Partnership (USAEP) for providing funding for and ongoing support to the *WISE* project;
- PDAMs Kabupaten Lampung Selatan, Kotamadya Bandung, Kotamadya Malang, and their respective local governments for their support in conceptualizing and implementing the CSS pilot surveys in January 2000;
- household and non-household respondents to the pilot survey as well as individuals who participated in pre-survey focus groups;
- various PDAMs in West Java that (i) provided the time of management and staff in early discussions of the CSS approach and (ii) participated in discussions of pilot CSS implementation in West Java;
- PDAMs from Sumatra, Sulawesi, Kalimantan, and Java that participated in and contributed to the CSS training in Batu Malang;
- participants in the CSS Stakeholder Workshop held in Jakarta in February 2000;
- Departamen Dalam Negeri for providing support to the *WISE* team and to the *WISE* project Steering Committee;
- PERPAMSI for providing the important time of a *WISE* project team member and for important guidance through the Steering Committee;
- IATPI for providing the important time of a *WISE* project team member;
- the USAID CLEAN-URBAN project for providing much of the initial "push" for starting the WISE project; and
- the RTI-Asia office for its continuous support to the *WISE* project.

1. Overview of CSS Implementation Manual

This *Implementation Manual* is intended to serve as a resource which can be used to help guide and support performance of a Customer Satisfaction Survey by any Perusahaan Daerah Air Minum (PDAM) in Indonesia.

As a manual focused on implementation, it is designed to provide users with an understanding of the tasks that must be carried out to *do* a Customer Satisfaction Survey and to begin to use the resulting data. To achieve these goals, the manual necessarily must address many topics. To address these topics and yet be as compact and easily usable as possible, it is written at several different levels of detail. It is—for topics that will be familiar to most readers—occasionally only an overview. For topics that are perhaps less familiar, it provides more detailed descriptions and instructions. For topics with which typical readers are unlikely to be familiar, it provides much more detailed background, descriptions, and instructions and can often function as a cookbook.

What is this manual *not*? Many things. It is not, for example, a textbook on the many possible tools for analyzing survey data. We do not discuss regression analysis, analysis of variance, factor analysis, or the many other techniques that are available to help understand survey data and the respective populations under study. This manual is also not a textbook on sampling. While we discuss the aspects of sampling that are directly relevant to PDAMs implementing a typical CSS, we do not cover the many possible approaches to sampling that are available. Nor do we discuss the statistical-mathematical basis of sampling. In a "Bibliography and Sources for Additional Reading" following the main text, we provide and describe several references which point the interested reader toward useful texts covering these areas.

The contents of the manual are organized as follows:

Section 2 contains a basic description of the process of carrying out a Customer Satisfaction Survey. It is an overview of the Customer Satisfaction Survey (CSS) process, designed to help the reader both to understand the process as a coherent whole. Section 2 also serves as an introduction to and preview of the following sections, which describe the component tasks of a CSS in more detail.

In **section 3**, we introduce two critical aspects of the survey process. First, we discuss the basic principles of sound questionnaire design—concentrating primarily on questionnaire *focus* and *ease of use*. Following this description, we focus more narrowly on what constitutes a good survey question and highlight some of the common difficulties and potential pitfalls of writing survey questions.

Section 4 presents an overview of survey field operations. In this section, we use the model of a typical day in a typical CSS to describe the management of the survey process. Section 4 focuses largely on the management of enumerators and questionnaires.

In **section 5**, we discuss at length several critical issues related to enumerators. Given the importance of enumerators in successful survey implementation, we spend significant time on the topics of selection, training, support, and monitoring enumerators.

Section 6 discusses sampling. It covers determination of required sample size, generation of a random sample, and stratified random sampling.

Section 7 describes handling the resulting survey data. The section covers both data entry and data cleaning, the final step prior to analysis.

Analysis of CSS data is taken up in **section 8**. Analysis is presented using data from a fictional PDAM. We trace the thought process of analysis, focusing on producing results on which the PDAM can act. Section 8 concludes with an outline of several components of an action plan derived from the survey data and analysis.

Several appendices contain materials which can be adapted by survey managers to help design, initiate, and manage the survey process:

Appendix 1 contains several sets of questions which can be used to conduct informal customer focus groups. These materials are appropriate for use during design of the survey's focus and *prior to* design of survey instruments.

Appendix 2 contains an enumerator training outline. This outline covers the topics that any enumerator training program must cover. The specific focus of individual surveys may require coverage of additional topics, but the topics shown in appendix 2 are necessary for all surveys.

Appendix 3 contains a more detailed and annotated version of the outline in appendix 2.

Appendix 4 contains a sample survey field log. This log, or an adaptation, is vital for tracking the survey process. Most importantly, the sample log shows a simple method of tracking completed questionnaires versus the sampling plan.

Appendix 5 contains versions of questionnaires used in the pilot survey. Specifically, appendix 5 contains questionnaires for (1) connected households, (2) non-connected households, (3) connected non-households, and (4) non-connected non-households. Additional questionnaires—perhaps focusing in a more detailed fashion on specific aspects of service or on narrower customer groups—can be designed using these questionnaires as a starting point.

Accompanying the manual is a **computer diskette** containing several spreadsheet and word processor files. These files are referenced throughout the manual and contain either examples of topics addressed in the manual or utilities designed to help PDAMs that are implementing Customer Satisfaction Surveys. Immediately below, we provide the name of each file and briefly describe its contents. (The word *language* used in italics indicates that there are two files—one in *English* and One in *Indonesian*. The first file name listed, for example, corresponds to the files "CSS Questionnaire Connected Household English.doc" and "CSS Questionnaire Connected Household Indonesian.doc"

- **"CSS Questionnaire Connected Household.doc"**
  This Microsoft Word file contains the questionnaire for connected households used in the *WISE* project's pilot surveys. (All questionnaires are in Microsoft Word.)

- **"CSS Questionnaire Connected non-Household.doc"**
  This file contains the connected non-household questionnaire used in the *WISE* project's pilot surveys.

- **"CSS Questionnaire non-Connected Household.doc"**
  This file contains the questionnaire for non-connected households used in the *WISE* project's pilot surveys.

- **"CSS Questionnaire non-Connected non-Household.doc"**
  This file contains the non-connected household questionnaire used in the *WISE* project's pilot surveys.

- **"Sample Size.xls"**
  This Microsoft Excel file shows how to determine the necessary sample size based on various desired levels of confidence and confidence intervals. (All ".xls" files are Microsoft Excel files.)

- **"Random Sample.xls"**
  This file contains a simple example of generation of a purely random sample using Excel's "Tools/Data Analysis/Sampling" feature.

- **"Data Entry.xls"**
  This file contains a basic data entry spreadsheet.

- **"Simple Analysis.xls"**
  This file provides and example of tabular and graphical analysis of a two-question survey.

This implementation manual as well as the accompanying files are also available for downloading from the internet at http://www.usaep.org/css/manual

2. Activities Required to Implement a CSS

Implementation of a Customer Satisfaction Survey consists of a number of related steps or activities. These steps proceed from the early stage of establishing survey focus and objectives to the later stages of internally utilizing information obtained from the survey and communicating the results of the survey to customers and other stakeholders.

The basic flow of the survey process is as follows:

- Design of survey objectives and focus
- Design, pre-testing, and revision of a survey instrument
- Design and selection of a sample

- Selection, training, and monitoring of enumerators
- Survey field operations
- Data cleaning
- Necessary re-sampling
- Data analysis
- Reporting/Communication

Each of these steps is an important part of the survey process. Each step should be planned and implemented with care. The success of later steps and of the whole survey process depends upon devoting the required attention to each activity. If activities are not well-designed and well-executed, it is possible that they will have to be repeated. Repeating activities is expensive, is a threat to the overall schedule, and can significantly impact the morale of participants.

In the following paragraphs, we briefly discuss each of these topics to give the reader an overall view of the survey process. The more complex and important steps are then discussed in more detail in sections 3 through 9.

survey focus

Determination of survey focus is a critical step in the survey process. In the context of Customer Satisfaction Surveys, most of the focus is, in a sense, already determined. Questionnaires will focus on various aspects of customer satisfaction with PDAM services, improvement priorities, and willingness to pay for improvements. But there are still decisions to be made:

- Will the survey focus on all connected customers, both households and non-households?
- Will non-customers be included?
- Will the focus be primarily on issues such as water quality, continuity, and pressure?
- Is understanding customer satisfaction with payment procedures, metering and other institutional features important for the PDAM?
- If so, will the focus of non-connected questionnaires be perceptions of PDAM service quality, willingness to pay connection fees, or both?

There is no blanket answer to these questions for all PDAMs. The answers depend on the situation and priorities of the PDAM and its customers. One PDAM might have acceptable service in terms of water schedule and pressure, but yet be seriously lacking in terms of payment points and customer service representatives. Another might be more concerned with serving its existing customer base than with expanding service. PDAM management and staff should devote significant time to thinking through these questions to design a survey that adequately addresses their particular situation and needs.

Most PDAMs are likely to already have some idea of their own priorities and the priorities of its customers. Internal meetings of PDAM management and staff to discuss and determine the PDAM's strengths, weaknesses, perceptions of customer attitudes, and improvement priorities can be a productive first step in the process of implementing a customer satisfaction survey. But the PDAM is of course only half of the picture. Customers are the other. Conducting focus groups and other formal and informal meetings with customer groups are absolutely essential steps in determining customer interests, needs, and priorities. Any and all contact with customers can be very useful in establishing the area on which survey instrument(s) and related activities should focus. In addition, such meetings

can represent important bridge-building steps toward the beginning of better understanding of and relationships with the PDAM's customer and stakeholder groups. Appendix 1 contains materials which can be used as a guide for conducting informal customer focus groups.

survey instrument(s)

The design, construction, and testing of any questionnaire is critical to the success of a survey. A well-thought-out questionnaire, which addresses the potential interests of different customer groups can produce very useful information. Information obtained from a survey instrument that does not deal with issues that concern customers, however, will be of little use to the PDAM. In addition, poorly-worded questions will be difficult if not impossible to analyze. Questionnaires should be pre-tested and revised prior to enumerator training. More revisions may be necessary following enumerator training. Survey instrument design is discussed in more detail in section 3, below.

sampling

The procedures for determining which customers will be interviewed is a critical step in the survey process. If (1) a sufficient number of individuals is interviewed and (2) these individuals are properly selected from the population, the results of the survey will be valid and we will be bale to make inferences about the population form the sample. Sampling is discussed at some length in section 6, below.

enumerators

Enumerators are the individuals who will implement the sampling plan and administer the questionnaire. Interviewers are the connection between the directors of the survey and respondents, and between data analysts and respondents. Properly selected and trained interviewers can make survey administration and analysis relatively smooth tasks. Improperly selected and/or trained interviewers can ruin a survey. The selection, training, and management of enumerators are discussed in section 5, below.

survey field operations

Survey field operations refers to all activities which take place while in the field. The many aspects of field operations—distribution of questionnaires to enumerators, monitoring of enumerators, proper record keeping, and other aspects—are discussed in section 4, below.

data cleaning

Data cleaning is the process of (1) checking completed questionnaires to ensure that questionnaires are sufficiently complete and that responses are within expected ranges and (2) making corrections where possible and discarding questionnaires/data that are not complete or difficult to interpret. Data cleaning is discussed in section 7, below.

re-sampling

Re-sampling refers to the process of conducting additional interviews to (1) replace questionnaires/data points that were not useful or (2) make adjustments for an incorrectly-implemented sampling plan. Re-sampling is discussed in section four, which covers survey field operations.

analysis
Analysis is the processing of raw data into usable information. Survey analysis typically results in the production of graphs, tables, and other outputs. These outputs communicate what can be learned from the survey. What are the main concerns and priorities of industrial customers? Are households in one kecamatan more concerned with water quality than households in other kecamatans? Are businesses satisfied with their communications and contact with PDAM management? Are households that are not satisfied with the quality of PDAM repair service willing to pay higher tariffs in exchange for more effective repairs? Analysis is discussed at length in section 8.

reporting
Reporting is the process of communicating the results of survey analysis to customers, non-customers, and other stakeholders. In addition to communicating results, a strategy of describing if and how results will be acted upon is critical as well. Carrying out a customer satisfaction survey creates the expectation among respondents that information gained from the survey will be acted on. Even a survey that is not well-publicized will quickly become common knowledge within the PDAM's service area. Strategies for reporting results are discussed in section 9 on use of survey data, below.

3. Survey Instruments

The survey instrument, or questionnaire, is the primary vehicle through which enumerators interact with interviewees. The questions within the questionnaire are ultimately the basis for analysis and reporting of results. Sound questionnaire and question design are, therefore, top priorities of the survey researcher. What constitutes good questionnaire and question design, however, is not always obvious. We are all used to asking people questions and thinking about their responses, but we typically do this in an informal setting where intuition can help us interpret ambiguous responses, or responses to vague or loosely-worded questions. By its very nature, however, a survey is a more demanding form of communication.

First, in section 3.1, we will discuss the principles of good questionnaire design and then, in section 3.2, we focus more narrowly on the design of good questions. While these points are closely related, we will disentangle them for the sake of clarity. Following these design discussions, in section 3.3 we will discuss the importance of the respondent identification number.

3.1 Questionnaire Design

A well-designed questionnaire will posses the following characteristics.

Questionnaires should focus only on topics that are necessary for the research which the survey is designed to support. While it is certainly tempting to obtain a wide range of information from a survey, it is best to address only topics that are of definite or highly-likely interest to the researcher(s). A technique that can help maintain a tightly-focused questionnaire is to work backwards from the outline of the analysis report and draft table shells to the questions required to

support the analysis and build the required tables. While analysis results almost invariably lead to the need to construct additional tables to help explain unforeseen features of the data, such an approach provides the discipline necessary to keep the size of the questionnaire in line with the size of the analysis task.

Questionnaires should be of a length and complexity that allows them to be administered in a reasonable amount of time. Obeying the previous rule on questionnaire focus will do the majority of what is required to keep questionnaire length and administration time at reasonable levels. However, even a questionnaire that focuses only on areas that are important to the researcher can be too long. Occasionally difficult choices must be made and questions that are relatively less important must be eliminated. The rules here are rather simple: (1) listen to the respondents during the pre-test, they will tell you if the questionnaire is too long, (2) discuss the issue of questionnaire length with enumerators during the enumerator training period, they also will typically not hesitate to say if the questionnaire is too long or too difficult.

Questionnaires should be easy to use. The layout of the questionnaire should be simple, the use of complicated skip patterns should be minimized, the font selected should be readable, it should be clear which responses go with which questions, and questions should be contained within a single page. These principles, while perhaps obvious, are nonetheless important. Questionnaires that are difficult to administer will result in enumerators and respondents that are confused and dissatisfied. This will result in skipped questions, incorrectly entered responses, and, ultimately, data that are of lower quality than is possible with a well-structured, clear questionnaire. Data cleaning will be more difficult and the results of the data analysis will be of lower quality. Again, listen to respondents and enumerators during training and pre-testing.

Questionnaires should contain an opening statement that provides the potential respondent with background information on the survey. Specifically, the opening statement should:

- clearly state the purpose and importance of the study,
- explain the importance of respondents and the information that they can provide,
- let the respondent know that his time is valuable and respected,
- address any privacy concerns that potential respondents might naturally have,
- describe the sampling process and how the respondent was selected for participation, and
- assure the respondent that there are no “right” or “wrong” responses and that all truthful and accurate responses are acceptable and valuable.

The following is the opening statement used for the connected household questionnaire in the first pilot test of the Customer Satisfaction Survey program in Indonesia.

Dear Respondent! My name is _____. I represent a PDAM/DDN water/wastewater study team. We are conducting a survey related to customer satisfaction with PDAM services and customer priorities for PDAM service improvements. This information will help us to design a program for guiding improvements in PDAM customer service for your PDAM and for PDAMs throughout Indonesia. From among Indonesia's 300 PDAMs, your PDAM has been chosen as a test case for this important work. From within this PDAM, your household has been selected at random to be interviewed. Information from households such as yours is vital in designing this program.

We assure you that your individual responses will not be disclosed to anyone. After questionnaires are completed, they will be processed by computer and no information on any single household will be disclosed to anyone. You can of course choose not to participate. The survey will take approximately 20 to 30 minutes and we will be asking you questions related to your household's:

- basic characteristics;
- sources and uses of water;
- satisfaction with PDAM services;
- priorities for improvements in PDAM services;
- contact and communication with your PDAM; and
- opinion of PDAM services relative to other service providers.

Are you willing to participate? *If NO, thank respondent and end interview.*

Thank you for participating. The usefulness of the results from this survey will depend on your sincerity and exactness in answering these questions. There are no "right" or "wrong" answers to *any* of these questions and you will not be judged in any way based on your responses. Please answer all questions as accurately and truthfully as possible. Again, thank you for taking the time to participate in this survey.

During training, enumerators should be instructed to read the opening statement—and all questions—*verbatim*. This practice helps ensure that all interviewees are responding to the same questions, based on the same information set. If we wish, as we almost always will, to state group averages for questions, it is crucial that respondents are answering identical questions based on similar background information. We will address these important points again in the sections covering question design and enumerator training.

Introductory questions should be used to help respondents (1) focus on particular, more detailed topics or (2) help respondents organize their thinking or recall past experiences. Let's begin with an example.

3.1 Does this household *ever* obtain water from the following sources?

	yes	no
3.1.1 PDAM dwelling connection	1	2
3.1.2 TA/HU/public tap	1	2
3.1.3 well	1	2
3.1.4 neighbors	1	2
3.1.5 purchased (not from PDAM)	1	2
3.1.6 lake, river, or other natural body of water	1	2
3.1.7 collection of rainwater	1	2
3.1.8 other	1	2

3.2 What is the *primary* source of water for this household?

PDAM dwelling connection	1
TA/HU/public tap	2
well	3
neighbors	4
purchased (not from PDAM)	5
lake, river, or other natural body of water	6
collection of rainwater	7
other	8

By asking respondents to recall from which sources they *ever* obtain water, they will be forced to mentally review their experience with these sources and will be better able to answer the question related to their primary water source. In fact, while question 3.1 was of some interest in this particular study, 3.2 was of more interest and 3.1 was included primarily to ensure better responses to 3.2.

Related questions should be located together, in thematically organized sections. Respondents are better able to answer questions accurately when they are in the "mind set" of thinking about a particular topic or area. See the questionnaires contained in Appendix 5 for examples of thematic grouping of questions. These categories represent fairly standard groupings for water sector questionnaires.

Skip patterns, also termed "filter" or "selection" questions, should be used to ensure that respondents answer only questions that are relevant to their situation. By forcing respondents to skip irrelevant questions, the time that is required to carry out an interview can be minimized. In addition, if interviewees respond to questions that are irrelevant to their circumstances, the responses. See the following example.

6.1.1 Is your dwelling connected to the PLN (electricity service)?
yes 1
no 2
don't know.....3

IF 6.1.1 IS 2, SKIP TO 6.2

6.1.2 Overall, how would you compare your satisfaction with PDAM service relative to PLN (electricity) service?
more satisfied with PDAM than with PLN.....1
equally satisfied 2
less satisfied with PDAM than with PLN.....3

Now that we have discussed the characteristics of good questionnaires, let's review some basic principles of good survey questions.

3.2 Question Design

First, let's define two important terms.

A **closed-ended** question is one for which a list of possible responses is provided. The interviewee selects one of the available responses. An **open-ended** question is a question for which the use is not provided with a list o possible responses, but rather provides their own response. Closed-ended questions are easier to analyze than open-ended questions. (The use of open-ended questions in focus groups is an excellent tool for designing closed-ended questions which will be used in questionnaires.)

Sound question design should obey the following principles:

Questions should focus on a single topic. Questions containing the words "and," "or," "both," or "either" should be reviewed carefully. It is possible that they will be difficult to interpret.

First, consider the following two questions from a household questionnaire used in a fictional PDAM in Kabupaten Sawah Indah.

3.4 Are you satisfied with the pressure or quality of the water from your PDAM connection?
yes 1
no..... 2

3.5 Would you be willing to pay a higher tariff in exchange for improvements in the pressure or quality of the water from your PDAM connection?
yes 1
no..... 2

This question focuses on two important aspects of PDAM service—the pressure and quality of water from a household's PDAM connection. A good PDAM customer satisfaction survey questionnaire would certainly address these issues. But a good question would not! A good question would address one of these issues. The primary problem with addressing two issues in one question is that interpretation of the results is difficult or impossible.

Suppose that 87% of the respondents in one service area reported that they are not satisfied (question 3.4) and 96% of the households that are not satisfied reported that they are willing to pay a higher tariff in exchange for improvements (question 3.5).

Suppose further that the PDAM decides to increase tariffs and use the proceeds to address the households' concerns. On what should the PDAM focus its attention and new resources? Should it attempt to fix leaking pipes and improve pressure, or should it invest in upgrading its treatment facilities? It cannot decide based on responses to questions 3.4 and 3.5.

Let's look first at question 3.4. It is possible that all 87% of the dissatisfied customers are dissatisfied with water pressure *only*. It is also possible that all are dissatisfied with water quality only. It is further possible that the dissatisfied 87% represent some combination of dissatisfaction with pressure and quality. We cannot determine what households are dissatisfied with on the basis of question 3.4.

Similar problems of interpretation apply to question 3.5 and they are, in fact, compounded by the poor design of question 3.5. A better design for 3.4 and 3.5 would be to separate the issues of water quality and pressure, with a satisfaction question for each and a willingness to pay question for each.

Questions should not lead the respondent, or otherwise suggest a correct response. Consider the following question.

> **4.8** Access to potable water is an important determinant of the health of children. Are you concerned with the important topic of water quality?
> yes 1
> no 2

Both the statement about children's health and the use of the word "important" are likely to lead many respondents who are not concerned with water quality to respond that they are. It suggests a social norm or generally accepted response. Respondents can easily be led into responses by the use of such words as "important" or "necessary." When writing opening statements or statements to introduce new sections of the questionnaire, one must also be careful not to lead or otherwise influence respondents.

For closed-ended questions, the responses provided should cover all reasonable potential responses from the interviewee. A question related to monthly PDAM bills, for example, should not have the following form.

7.9 On average, what is your total monthly bill from your PDAM?
Rp 5.000 — Rp 10.000 1
Rp 10.001 — Rp 20.000 2
Rp 20.001 — Rp 30.000 3
Rp 30.001 — Rp 40.000 4

Why is this a poorly designed question? It ignores many possible—and not at all unreasonable—values. It is likely that there are households whose monthly PDAM bill averages less than Rp 5.000. It is also possible that some households have monthly bills in excess of Rp 40.000. Since the range of potential responses is often difficult to know prior to administration of the questionnaire, closed-ended questions should contain responses which span a wide range of values.

Questions should use simple wording and should avoid the use of uncommon or difficult to understand words or phrases. Complicated wording is difficult for both enumerators and for respondents. Also, respondents will be reluctant to ask enumerators the meaning of unfamiliar terms. Respondents who do not understand particular terms will nevertheless tend to answer to the questions, despite not fully understanding their meaning. Many respondents will effectively be providing random answers. When we carry out analysis of such responses, the true tendencies of respondents will not emerge. Consider the following example.

3.15 Although it has been demonstrated that on average turbidity tends to be higher during the rainy season, it can also be excessive during the dry season.
Is the water from your household's PDAM connection ever excessively turbid during the dry season?
yes 1
no 2

While this question has several design problems, the most obvious are its complicated structure and use of the technical term "turbidity." Enumerators will have difficult time reading the question and respondents will have difficult time understanding the question and responding to the question. Many respondents, however, will randomly select a "yes" or "no" response. In reality, water in a particular area might almost never be excessively turbid during the dry season and, say, 93% of the population might respond "no" *if they understood the question*. If a sufficient number of respondents are providing random answers—with a 50% probability of responding either "yes" or "no"—analysis of the data will not reveal the true situation in the area. Results will be biased toward "yes" responses by the poor design of the question. While this is a very serious problem, there is yet another discussed in the following paragraph.

In section 5, below, we will discuss enumerator behavior. One aspect of enumerator behavior is particularly important and relevant to question 3.15, above. Many if not most enumerators will behave as effort-minimizers. They want to do as many interviews as possible in as short a time as

possible. The frequent practice of payment on a per-completed questionnaire basis reinforces this tendency. Despite being trained to read questions *verbatim*, in the face of such a poor question, enumerators will not do so. The likely (even logical!) behavioral response of an effort-minimizing enumerator will be to attempt to simplify this question. This simplification will likely take the form of shortening the question and/or changing its wording. Either of these actions will mean that different enumerators will be asking slightly different questions. When interviewees are responding to different questions, we cannot aggregate their responses and report group averages or the proportion of respondents providing various responses. So, not only are such questions difficult to administer, they will tend to result in data that are biased and impossible to analyze.

Poorly designed, overly complicated questions are bad if enumerators read them verbatim and bad if enumerators adapt their wording. Both cases are a potential disaster for survey analysis!

Questions should not deal with topics that are excessively sensitive for potential respondents. Obviously there are occasions and surveys where sensitive issues must be addressed—a health-related survey is one example. A good rule to follow is that the most sensitive topics and questions should occur *as late as possible* in the questionnaire. If the respondent refuses to continue the interview following sensitive questions, by placing them near the end of the questionnaire, we can possibly obtain much useful information before the interview is terminated. Within the context of Customer Satisfaction Survey in the water sector, it is difficult to imagine that sensitive areas will be covered. The one exception is questions related to household income and expenditure. Generally, these questions are placed in the questionnaire's final section.

3.4 Respondent Identification Number

For several important reasons, it is vital that we are able to identify each respondent within a survey data set. First, we review these reasons and then we discuss how the information required for identification numbers is collected and transformed into a unique number for each respondent.

A unique respondent identification number allows survey managers and analysts to:

- **check or "clean" data**
  As we will discuss in section 7, below, a vital step in the survey process is cleaning the survey data set. Incomplete identification numbers must be corrected, or respondents removed from the data set. Responses that fall outside of acceptable ranges must be checked versus the physical questionnaire. In cases where responses were entered into the data file incorrectly, the are re-entered correctly. Where responses fall outside of the allowable range, they must be noted and excluded from analysis.

- **group respondents according to different characteristics**
  It is frequently important to perform analysis for subgroups of respondents. The most common case will be where PDAM management is interested in carrying out analysis for one wilayah, kecamatan, kelurahan, service area, or other geographic unit. Understanding the different concerns and priorities of a PDAM's different geographic areas can be a valuable tool in targeting services, service improvements, and service expansion.

How do we obtain the information required to build a unique respondent identification number? What information do we need? Let's look at an example from the pilot surveys in Lampung Selatan, Bandung, and Malang. Appendix 5 contains the questionnaires used for connected and non-connected households and non-households. The cover page of a questionnaire is always where the information required to build an identification number is collected. Immediately below, we reproduce a portion of that cover page for the connected household questionnaire.

Interview Data

PDAM:	**A**	_____
Location 1:	**B1**	_____
Location 2:	**B2**	_____
Interviewer ID number:	**C**	_____
Date of interview:	**D**	_____ ***month*** **E** _____ ***day***
Interview number:	**F**	_____

HOUSEHOLD IDENTIFICATION NUMBER:

A	**B1**	**B2**	**C1**	**C2**	**D1**	**D2**	**E1**	**E2**	**F1**	**F2**

After determining if a potential respondent is willing to participate in the survey, the interviewer must record several pieces of information related to the interview. This information is used to construct the household identification number. The pieces of information are structured such that no two households can have the same identification number. Let's review these pieces of information. (Note that we will continue to use the example of a household, but that the same rules hold for non-households as well. Also note that interviewers fill out only the top portion of this page and the household identification number is built in the table at the bottom of the page by data entry personnel or survey managers.)

- **PDAM**
  Within the pilot survey phase of the Customer Satisfaction Surveys project, the survey was carried out in three PDAMs. For this reason, we needed a code to distinguish each PDAM. This portion of the identification number will not be included in a *single* PDAM's identification codes.

- **Location 1**
  Typically, surveys are implemented in multiple areas within a PDAM. In Bandung, for example, the pilot survey was carried out in all wilayahs (location 1) and in multiple kecamatans (location 2) within each wilayah. In order for analysis to be carried out for individual wilayahs and/or kecamatans, analysts need a code which can be used to identify and select each area. The location 1 and location 2 codes provide this information. In the case of Bandung, location 1 was a pre-

assigned code for each wilayah and location 2 was a pre-assigned code for each kecamatan within each wilayah.

- **Location 2**
  See "location 1" explanation.

- **Interviewer ID Number**
  It is also important to be able to identify which enumerators performed each interview. In some cases, interviewers might not carry out interviews properly and the questionnaires for these interviewers must be discarded. While this is unfortunate and can result in additional effort for managers, analysts, and other interviewers, the interests of having correct survey data override those of minimizing survey effort. Each interviewer should be assigned a unique number during training. *If interviewers are replaced at any point during administration of the survey, their number must not be assigned to their replacement. The replacement must be assigned a new, unique interviewer ID number.*

To this point, the household identification number shows where interviews occurred and which interview carried out the interview. Multiple interviews will be carried out by each interviewer within each survey area, however. For households to be uniquely identified, we therefore also require information on when the interview was carried out. This information is obtained using the following two items.

- **Date of Interview**
  This portion of the household identification number provides the date on which the interview occurred.

- **Interview Number**
  The interview number is the number of the interview done on that day by each individual interviewer.

Together, these pieces of information uniquely identify each household, non-household, individual, or other sampling unit. The exact use of the identification number will be discussed in detail in the sections on data cleaning and data analysis, below.

4. Survey Field Operations

The term **survey field operations** refers to all day-to-day activities required to implement a survey as well as the management of these activities. While simple, survey field operations are important and should be carried out in a careful and methodical manner. Most PDAM managers and many PDAM staff members will be familiar with the activities and work modes that comprise a survey—working in teams, managing complex tasks, and overseeing non-PDAM employees, for example. The activities required to manage a survey will, therefore, be different from many of their past activities only in terms of particulars.

In this section, we briefly describe management of survey field operations from the perspective of a "typical" day in a typical survey project. Unexpected events, requiring departures from the planned schedule, will occur in all survey projects. Few days will be purely typical. For this and other reasons, we recommend that more than one PDAM staff member be assigned to management of the survey process. Both a designated survey manager and a potential replacement should be familiar with all survey instruments, the sampling plan, the plan for field operations, and other aspects of the survey. When the manager is called away to address survey-related problems, the replacement can step in with little disruption to the day's activities.

A typical survey day can be broken down many ways. We use a simple division that bundles the various activities which will be carried out in a typical day. This division has 5 parts:

- activities which take place prior to the morning enumerator meeting;
- the morning meeting with enumerators;
- monitoring/support of enumerators;
- the evening meeting with enumerators; and
- post-evening meeting activities.

In the following sections, we discuss each of these activities in a bit more detail.

4.1 Pre Morning Meeting with Enumerators

Before the morning meeting with the enumerator team, the survey manager should:

- ensure that all questionnaires from the previous day have been reviewed and logged;
- review all enumerator location assignments; and
- review plans for the morning enumerator meeting.

It is critical that the survey manager is thoroughly prepared for the morning enumerator meeting. Being fully prepared will help ensure that enumerators are deployed properly, that they have the information they need to carry out their assignments, and that they have confidence in the survey manager. A sense of confidence in the survey manager is critical. When enumerators are not confident in the survey manager, they will not see the benefit of asking the manager questions they might have, they will be less likely to correctly implement the sampling plan, they will be much less likely to share interesting (and important) information which they learn in the field, and they will not treat the survey as a serious undertaking.

4.2 Morning Meeting with Enumerators

The morning meeting is designed to accomplish several important goals. The main focus is discussing planned activities for the current day. It is important also, however, to discuss any relevant issues related to the preceding survey day(s). As with all meetings, enumerators should be encouraged to ask questions and make suggestions at any point. The following plan is a good starting point for organizing a typical morning meeting.

First, **questionnaires not received during the previous evening's meeting must be collected, reviewed, and logged**. This activity, carried out prior to the planned discussion with enumerators, will allow the survey manager to make any final revisions to the current day's enumerator location assignments.

Second, it is important to **briefly review the previous days' activities**. In particular, the survey manager should describe progress in implementing the overall sampling plan. This will allow enumerators to understand the role of planned activities for the coming day in fulfilling the overall objectives of the survey.

Third, the survey manager should **review any issues or problems which arose during the preceding survey day and will affect, or are likely to affect, the current day's activities**. Specifically, the survey manager should discuss strategies that have been designed and/or adjustments that have been made to handle these problems.

Fourth, **daily enumerator location assignments should be given and reviewed**. Enumerators should be clearly instructed where they are to conduct interviews as well as how many questionnaires of each type they are expected to administer in each area. We recommend the use of daily, enumerator-specific maps for these assignments.

After enumerator location assignments have been reviewed, the **survey manager should lead a question and answer session**. Questions concerning the day's sampling plan should be given particular attention.

Finally, the survey manager should announce **(1) which enumerator(s) he/she will accompany to the field** during the current day and **(2) which questionnaires from the previous day(s) will be followed-up** with calls and/or visits from survey management.

4.3 Intra-day Monitoring and Support of Enumerators

support

The survey manager should be available to the enumerators at all times during the survey day. If the survey manager does not have a hand phone, one should be provided for his/her use during the survey period. Enumerators should be given this number and encouraged to call with any questions, regardless of how small.

monitoring

The survey manager should accompany at least one enumerator into the field on each day. This enumerator should be selected at random at the conclusion of each morning enumerator meeting. Actually drawing names out of a hat can be a useful way to ensure that enumerators (1) understand that they can be selected on any given day and (2) do not feel that selection is determined by manager favoritism. If it is possible to use additional PDAM staff, multiple enumerators can be monitored on certain days. A heavier-than-announced level of monitoring is ideal early in the survey process.

4.4 Evening Meeting with Enumerators

A set time should be established for an evening meeting of the survey team. As with the morning meeting, the evening session is designed to accomplish several goals. These goals center primarily on receiving and checking questionnaires, but issues that arise during the day should be discussed. The following outline describes a full evening enumerator meeting.

First, the survey manager should **receive the day's completed questionnaires from enumerators**. The process of receiving includes review of questionnaires, particularly the respondent identification number for each; making corrections as necessary to identification numbers; logging of questionnaires; and payment for properly completed questionnaires.

Second, the survey manager should lead a group discussion of other **issues, problems, or which arise should be discussed**. As with the morning meeting, the survey manager should encourage all enumerators to discuss any issues which they feel are interesting and/or important. Problems for which the survey manager does not have an immediate solution can be revisited during the following morning's meeting.

All enumerators should be instructed to attend each morning's team meeting. For enumerators who absolutely cannot attend the following morning's meeting, **location assignments for the following day can be given**. They should be instructed, however, to contact the survey manager the following morning before proceeding to the field to learn if adjustments have been made to their assignment.

4.5 Post Evening Meeting Activities

The survey manager's day does not end with the conclusion of the evening meeting with enumerators. Following completion of the evening, the survey manager has several activities which must be carried out.

First, the manager must **store completed questionnaires**. Lost questionnaires are costly to replace and can result in serious problems with implementation of the sampling plan. A systematic filing scheme should be used so that *all* individual questionnaires can be easily accessed and located by identification number.

In the evening, the survey manager must **make a full enumerator deployment plan for the following day**. After the number of completed questionnaires for each survey area has been tabulated, the survey manager should calculate the number that must be completed the following day. This calculation will be based on the number of available enumerators, assignments made during the evening enumerator meeting, the total required sample for each area, and the number of completed questionnaires for each area.

To the extent that time is available and to the extent that the survey manager also has a data entry/analysis role, **evenings can be used for data entry, data cleaning, and preliminary analysis**.

5. Enumerators

Enumerators are second only to respondents in terms of importance within a survey project. With a well-selected, -trained, and –monitored group of enumerators, administration of a survey can be a smooth and relatively problem-free process. When these conditions are not met: survey management can be difficult; the survey process can be awkward and even confrontational for respondents, enumerators, and managers; and the data that result can be almost useless. The importance of enumerators cannot be overstated.

Enumerators:

- represent the primary link between survey designers/analysts and the population under investigation,
- are responsible for administering the questionnaire,
- are responsible for implementing the sampling plan, and
- are a vital source of information and insight for survey designers, managers, and analysts.

The current section discusses enumerators from several vantage points—proper selection, training, monitoring, and management. We also briefly discuss some important points related to enumerator behavior and compensation.

5.1 Enumerator Selection

Enumerator selection is ideally a two-step process. Potential enumerators should be informed *when first contacted* that the following is the process which will be followed. Merely stating that there will be interviews for the positions will likely do a great deal of the selection work. First, potential interviewers should be screened in one-on-one interviews prior to training. Talk with the applicants, have them read and interpret questions from the questionnaire, quiz them on their knowledge of the survey location(s), and observe their appearance and behavior. In practice, such interviews should eliminate all or almost all interviewers that do not meet the desired enumerator profile. The second step is observation of enumerators during training.

(If possible and affordable, it is best to hire enumerators that are not PDAM employees, or to send PDAM employees into the field dressed in "civilian" clothes, not in their PDAM uniform. It is possible, even likely, that respondents will not be 100% truthful if they perceive that their responses might offend the enumerator. Worse yet, respondents might fear repercussions from responses that are critical of the PDAM's service or management.)

What is the profile a good enumerator? Enumerators should ideally possess all of the following characteristics.

Enumerators must be sufficiently literate and numerate to administer the questionnaire, respond to interviewee questions, and administer the sampling plan. A questionnaire must be read to respondents *verbatim*, so literate enumerators are obviously necessary. Enumerators can be screened either (i) in interviews prior to training or (ii) during the training

process, particularly as mock interviews are carried out. When possible, eliminate unsuitable interviewers prior to training, but do not hesitate to dismiss applicants following training if they have serious difficulties reading or interpreting the meaning of questions. While it may be awkward to dismiss enumerators, the awkwardness pales in comparison to the difficulties of correcting later problems caused by the use of inappropriate enumerators. In addition, most surveys deal at least in part with information related to prices, schedules, and other numerical areas, so numeracy is a vital interviewer characteristic as well. The ability to read and use maps to locate survey areas and assignments is almost always necessary as well.

Enumerators must be acceptable to the population within which they will be conducting interviews. If enumerators are not acceptable to the target population, they will frequently not be able to gain access to potential respondents. If, for example, one of the aims of the study is to interview women, in more conservative rural areas it might be necessary to use exclusively female enumerators. If senior managers or engineers in firms are to be interviewed, shabbily dressed enumerators will not be taken seriously and will often be turned away. These factors can lead to higher refusal rates, incomplete interviews, higher survey costs, and lower quality data. Choose appropriate enumerators and insist that they appear and act in a professional manner when in the field. Request that enumerators arrive at training dressed professionally and during training do not hesitate to point out what constitutes acceptable dress and behavior.

Enumerators should be residents of the area within which they will be conducting interviews. Familiarity with the survey area is important primarily for implementation of the sampling plan. Enumerators must be able to find the areas in which they will work. Frequently, assignments will be given to enumerators based on particular streets, buildings, or neighborhoods. Enumerators who have the ability to accurately and quickly locate and move around within these areas will help achieve the objectives of the survey within the budgeted amount of time.

Enumerators must be available during the time in which the survey will be conducted. While seemingly obvious, this is an important point. When enumerators fail to show up on a given day, the survey schedule is disrupted. When enumerators frequently do not show up or quit altogether, additional enumerators must be located, screened, and trained. This can strain both the personnel and financial resources allocated to the survey project.

The experiences with enumerator selection and use from the pilot surveys in Lampung Selatan, Bandung, and Malang are instructive and interesting.

Lampung Selatan

In Kabupaten Lampung Selatan, enumerators were drawn exclusively from Karangtaruna, a kelurahan-level youth association. Enumerators were between 18- and 25-years-old, with education ranging from complete high school to complete university. Four of the ten enumerators were female. There were some problems with access to different areas for female enumerators. None had motorcycles and were therefore unable to carry out interviews in more rural areas. Female enumerators were, however, able to gain access to households without problems. Female enumerators completed far fewer questionnaires per day than did their male counterparts. Respondents were more interested in discussing PDAM service with female enumerators than with male enumerators. In addition to their value as enumerators, we particularly recommend the use of female survey staff in focus groups and other, more discussion-oriented CSS activities. Enumerators, irregardless of gender,

experienced some instances of respondent hostility. These incidents occurred only with respondents who were very dissatisfied with PDAM services.

Bandung

In Kota Bandung, enumerators were selected largely from ITB, the Bandung Institute of Technology. Obviously these individuals were sufficiently literate and numerate to carry out their assignments effectively. Also, many were studying sanitary engineering or related fields and were quite familiar with water-related issues. All were either from Bandung or were long-term residents of Bandung and were therefore sufficiently familiar with Bandung to implement the sampling plan. At the outset of the study, enumerator name tags were given to each of the enumerators. These nametags showed clearly that the enumerators were members of a water sector study team sponsored by various stakeholder organizations. It quickly became obvious that household and non-household interviewees were more likely to be cooperative and responsive when the interviewers stated that they were students at ITB. Interviewers who were students at ITB therefore adopted this approach.

Malang

The experience with enumerators in Kota Malang was generally excellent. Enumerators were drawn from Brawijaya University, a local administration academy, a local technical academy, and a smaller number from a local senior high school. Enumerator backgrounds included study of civil engineering, economics, and law. All were able to easily implement the sampling plan. The only difficulty reported—initial difficulty understanding some questions within the survey—was with several enumerators from senior high school.

5.2 Enumerator Training

As explained at the outset of section 5, enumerator training is critical to the success of any survey project. Well-trained enumerators will:

- understand and be able to communicate the objectives of the study,
- understand the content and structure of the survey instruments being used within the study,
- know the importance of treating respondents with respect,
- be able to implement the sampling plan established by the study's designers and managers,
- understand the importance of following required survey operations procedures, and
- represent an important channel through which survey analysts and managers can better understand the population that they are studying.

5.2.1 Outline of Enumerator Training

To effectively train enumerators such that they have these characteristics when they take to the field, the following topics should be covered. (Note that an annotated version of this outline, complete with detailed notes for trainers, is provided in Appendix 3.)

- **Study and Training Objectives**
  Purpose of CSS study
  Purpose of enumerator training

Importance of enumerators
Importance of interviewees

- **Review of Questionnaires**
  Interview data
  Opening statement
  Bias
  Simulated interview

- **Review of Sampling Strategy**
  Location assignments
  Household selection

- **Questionnaire Handling**
  General issues
  Frequency of return to study director

- **Problem Interviews**
  Interrupted interviews
  Aggressive interviewees

5.3 Enumerator Monitoring

Monitoring of enumerators is a central aspect of survey management. While monitoring is an important way to detect, understand, and deal with problems, it is also a vital form of support to enumerators. The plan for enumerator monitoring should therefore be described to enumerators as both support and monitoring. Enumerators must feel that survey managers are there to help them do their jobs, but they must also understand that their performance will be checked and evaluated.

Survey managers must ensure that:

- enumerators have the resources—sufficient training, local knowledge, instructions, maps, and transportation—to correctly implement the sampling plan,
- enumerators are carrying out interviews properly,
- questionnaires are being returned in a timely manner, and
- all interview data are entered promptly, correctly, and fully on all questionnaires.

These objectives can easily be achieved by the use of a three part interviewer monitoring strategy.

direct monitoring

Survey managers should spend time in the field with each enumerator. There is no substitute for occasional direct observation of enumerators. Do enumerators truly understand the importance of their role in executing the sampling plan? Are enumerators reading opening statements and questions *verbatim*? Are enumerators comfortable with the contents of the questionnaire? Are they recording interview data immediately following each interview? At the outset of interviews, the survey manager should inform the respondent that the manager might interrupt the interview to provide comments,

suggestions, or other support to the enumerator. Do this in a friendly manner that focuses on the support role of the manager. All enumerators should be clearly told during training that survey managers will select enumerators during the early stages of the survey and accompany them on several interviews.

follow-up monitoring

If the nature of the survey and the survey environment make it is possible to obtain addresses and/or phone numbers of respondents, enumerators should be told that survey managers will make follow-up telephone calls or visits to a random sample of each enumerator's interviewees. While there are no hard-and-fast rules for the frequency of follow-up calls/visits, following-up at least one of each enumerator's daily interviews would be ideal, resources permitting. Such calls or visits are intended to (1) ensure that interviews are actually occurring, (2) learn if enumerators are behaving in a friendly and professional manner with respondents, and (3) learn if there are problems or other issues that respondents wish to discuss.

checking of questionnaires

Ideally, enumerators should return completed questionnaires at the conclusion of each day. If daily return is not possible, enumerators should spend no more than two days in the field before returning questionnaires. When questionnaires are returned, survey managers should (1) check the interview data on each questionnaire, (2) review questionnaires for completeness, (3) discuss any problems that the interviewers have encountered in the field, and (4) discuss any interesting comments made by respondents.

By following this three-part monitoring and support strategy, enumerators will feel confident that they are being supported and will understand that their actions are being evaluated.

5.4 Enumerator Behavior and Compensation

Enumerators are like any other group of employees. Many will be devoted to and interested in their work and will carry out their tasks in a professional, energetic manner with a minimum of oversight. Others will seek to do the minimum amount of work possible, or even to avoid doing assigned work—yet still be paid. Proper selection and training of enumerators following the principles described above can result in a group that is comprised largely of the first type, or at least can serve to minimize the number of the second type. Survey managers should devote serious attention and energy to selection and training. Enumerators should nevertheless be monitored assuming that some effort-minimizers will slip through the selection process.

Enumerators can be compensated in different ways, or need not be compensated at all. If PDAM employees are used as enumerators, additional compensation will not be necessary. If, however, enumerators are hired from outside of the PDAM, they will need to be compensated.

Enumerators can be compensated on an hourly or daily wage basis. Such compensation, however, introduces no incentives for enumerators to complete more interviews within a given day. In fact, a wage system can introduce incentives for enumerators to complete fewer questionnaires, lengthening the period of the study and increasing their earnings. A wage system of compensation is typically only used in longer-term studies and in situations where enumerators can be more closely monitored.

(Of course, enumerators could be paid on a wage basis, with a minimum number of questionnaires required per day, but this is really equivalent to the following compensation scheme.)

In practice, the most common way in which enumerators are compensated is on a per-completed-questionnaire basis. Such a compensation scheme rewards enumerators for doing a larger number of interviews per day and can be a more efficient way to use project resources. This incentive is good, up to a point. Payment on a per-completed-questionnaire basis can induce dishonest enumerators to fill out questionnaires on their own, without actually conducting interviews. This possibility is yet another reason to treat enumerator monitoring and questionnaire follow-up seriously.

6. Sampling

6.1 Brief Overview of Sampling Concepts

Sampling, the process of selecting individuals, households or non-households to be interviewed, is critical to the success of any survey. Sampling is also a topic which is difficult for many people to understand. For these reasons, we discuss sampling at greater length than any other topic in the *Customer Satisfaction Survey Implementation Manual*. This is not a textbook on sampling, however, so we do not present the mathematical basis for sampling techniques nor do we describe the myriad sampling techniques which can be used to respond to different objectives, conditions, and resource constraints. For discussions of the mathematics of sampling, the interested reader can consult any advanced statistics text or specialized text on sampling. We cite several useful resources in the bibliography.

When good data on a population of interest are available, sampling can be carried out according to well-established statistical principles. Such samples—which are obtained using what is either termed "formal" or "probability" sampling—can be used to carry out analysis from which confident statements about the characteristics of the population can be made. Fortunately, it should typically be possible to sample the customers of a PDAM according to these principles. PDAM's are, of course, businesses and as businesses they keep and maintain records of their customers to facilitate billing and other administrative matters. These lists, if complete and up-to-date, are an ideal base for forming a sound sample of the PDAM's customers. Utility customers are typically thought of as one of the easiest groups for sampling due to the importance of customer lists to the utility.

Before we proceed, let's pause and define a few terms that are necessary to discuss surveys and sampling.

- A **survey** is some means of obtaining information about the characteristics of a group by observing members of the group.
- **Population** is a term that denotes a group of interest or focus within a survey.
- A **sample** is a subset of a population.
- **Sampling** refers to the process of selecting cases from a population that will be interviewed within a survey.
- A **sampling frame** is a list of members of a population which can be sampled.
- **Sample size** refers to the number of cases interviewed within a survey.

- A **population characteristic** is a measurement of a characteristic of a population.
- A **sample characteristic** is a measurement of a characteristic of a sample.

These terms, their meanings, their uses, and their importance will become clear as we use them in the discussions and in the examples which follow. Please refer to these definitions as needed when reading the text below. We will define other terms as we proceed.

For PDAMs, the two populations that are typically of most interest are (1) the population of connected household customers and (2) the population of non-household connected customers—for example businesses, enterprises, and social institutions. It is from these populations that PDAMs will most frequently construct samples and conduct analysis. PDAMs might also have reasons to sample non-connected populations, for example if they wish to investigate issues related to connection fees, or the perceptions of non-customers related to PDAM services.

The fundamental goal of sampling is to select a subset of the population, a sample, that is suitable for analysis. A sample is suitable if it (1) is of sufficient size and (2) is selected from the population such that the risk of obtaining results from analysis of the sample that do not accurately reflect the characteristics of the population is within an acceptable range.

Several aspects of sampling require further discussion. This discussion is contained in sections 6.2 and 6.3, below.

6.2 Sample Size

While it is possible to interview, for example, all households that are customers of a PDAM, it is generally neither feasible nor necessary to do so. (Interviewing all customers is termed a customer **census**, not a customer survey.) There are two principal reasons that sampling is typically preferred to census taking. These reasons center on issues of *relative costs* and *necessity*.

If non-employee enumerators will be paid for carrying out interviews, the monetary costs to the PDAM of conducting a customer census typically would be many times the cost if only a sample of households is used. In cases where PDAM employees will be used as enumerators, the opportunity costs of their time would be high—they would be away from their normal jobs and performance in other areas would suffer, or the completion of routine tasks would be delayed.

In addition to cost-based arguments against census sampling, it is simply not necessary to interview all customers in order to obtain useful and accurate information. Let's consider some examples.

If we survey 10 households from a population of 10 households and estimate a population characteristic—average height for example—we are 100% certain that our results do not differ from those of the population. We are 100% certain because our estimate is based on the entire population. We have measured a population characteristic directly. We are, therefore, confident that our results accurately represent the population we are trying to study.

If we survey 8 households, however, we are somewhat less confident that our results accurately represent the population we are trying to study, because they are obtained from a sample of the

population. We are less confident because our estimates are measured sample characteristics, not population characteristics. It is possible that the two individuals whose height we did not measure are the tallest within the population. In this case, our estimate of the average height within the population could significantly understate the true average height of the population. If we survey only 5 households—50% of the population—we are even less confident that our results will accurately represent the population we are trying to study.

For any given population size, increasing the size of a sample increases our confidence that the characteristics of the sample are similar to those of the population.

The sample size that is selected is generally the smallest sample which satisfies our desire for certainty about results. This level of certainty is typically expressed in a two part manner. For example, we might read a survey analysis report which states that a sample size was chosen such that the researcher is "… 95% confident that the mean height for all men in Sulawesi is within ±3% of the mean for the sample." Let's examine both parts of this statement.

The second portion of the statement, "… within ±3% …" is a percentage above and below estimated population characteristics. This percentage above or below is termed a **confidence interval**. The confidence interval specifies how precise an estimate the researcher requires. A smaller confidence interval requires a larger sample size.

The first portion of the statement "… 95% confident …" refers to the researcher's tolerance for the risk of error, or the risk that his estimate of the average height for men in Sulawesi is outside of the stated confidence interval. Thus percentage is termed **level of confidence**. A higher level of confidence, all other things equal, requires a larger sample size.

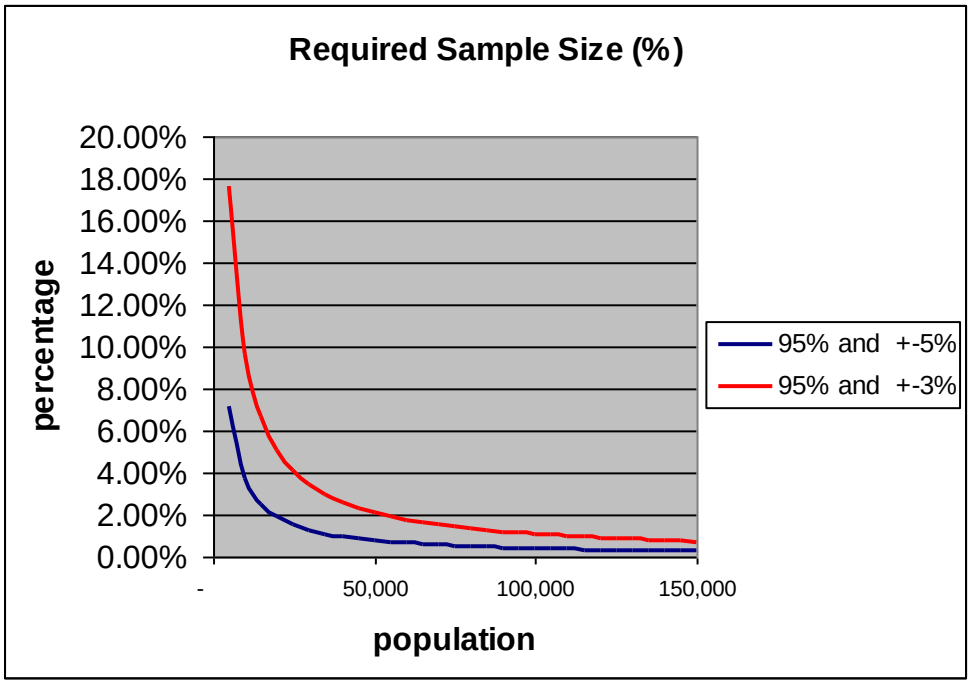
Let's illustrate these points with two fictional examples.

PDAM Kotamadya Havana, with exactly 100,000 household connections, wants to know the percentage of its household customers that is willing to pay a higher tariff in exchange for increasing the number of payment points. The management of the PDAM has decided that, if more than 50% of its household customers are willing to pay a higher tariff, it will increase the number of payment points. In order to estimate this percentage ±5% and with 95% confidence, PDAM Havana must interview (or sample) 383 connected households, or 0.38% of all connected households. This figure is taken from table 6.2.2, below. A smaller sample will yield results that are not within this margin of error (±5%) at this specified level of confidence (95%).

What about a nearby PDAM, PDAM Kabupaten Miami, with 5,000 connected household customers? To obtain the same level of confidence, PDAM Kabupaten Miami must sample 357 connected households, or 7.14% of the population. This shows an important and interesting fact: *as population size increases, we can survey a smaller percentage of the sample to obtain a given confidence interval and level of confidence*. (The reader is referred to any intermediate-advanced statistics or sampling text for a rigorous, mathematical explanation of why this is true.)

The graph below show the relationship between (1) population size and (2) required percentage of population to be sampled. The graph is based on two levels of survey certainty. The lower (blue) curve shows the correspondence between population size and required sample size for a 95% level of

confidence and a confidence interval of ± 5%. The higher (red) curve shows a 95% level of confidence and confidence interval of ±3%. As these two curves illustrate, for any given population size, a higher level of confidence and/or a tighter confidence interval requires sampling a higher percentage of the population.



The following tables contain the required sample sizes and percentages for various population sizes, levels of confidence, and confidence intervals. Note that percentages have been rounded to two decimal places. Included within the "Customer Satisfaction Survey Implementation Manual" diskette are two versions of a spreadsheet file ("Sample.xls" and "Sample.wk1") which can be used to determine the required sample size for any population size.

Table 6.2.1 Required Sample Sizes: 95% level of confidence

	95% level of confidence		
population	**± 3%**	**± 5%**	**± 10%**
500	250	218	81
1,000	500	278	88
1,500	624	306	91
2,000	696	323	92
3,000	788	341	94
5,000	880	357	95
10,000	965	370	96
20,000	1,014	377	96
50,000	1,045	382	96
100,000	1,058	383	96

Adapted from Rea and Parker (1997).

Table 6.2.2 Required Sample as % of Population: 95% level of confidence

	95% level of confidence		
population	**± 3%**	**± 5%**	**± 10%**
500	50.00%	43.60%	16.20%
1,000	50.00%	27.80%	8.80%
1,500	41.60%	20.40%	6.07%
2,000	34.80%	16.15%	4.60%
3,000	26.27%	11.37%	3.13%
5,000	17.60%	7.14%	1.90%
10,000	9.65%	3.70%	0.96%
20,000	5.07%	1.89%	0.48%
50,000	2.09%	0.76%	0.19%
100,000	1.06%	0.38%	0.10%

Adapted from Rea and Parker (1997).

The user will observe the large decrease in the percentage of the population that must be sampled as the size of the population increases. Tables 6.2.3 and 6.2.4 contain required sample sizes and percentages, respectively, at the 99% level of confidence.

Table 6.2.3 Required Sample Sizes: 99% level of confidence

	99% level of confidence		
population	**± 3%**	**± 5%**	**± 10%**
500	250	250	125
1,000	500	399	143
1,500	750	460	150
2,000	959	498	154
3,000	1,142	544	158
5,000	1,347	586	161
10,000	1,556	622	164
20,000	1,687	642	165
50,000	1,777	655	166
100,000	1,809	659	166

Adapted from Rea and Parker (1997)

Table 6.2.4 Required Sample Sizes: 99% level of confidence

	99% level of confidence		
population	**± 3%**	**± 5%**	**± 10%**
500	50.00%	50.00%	25.00%
1,000	50.00%	39.90%	14.30%
1,500	50.00%	30.67%	10.00%
2,000	47.95%	24.90%	7.70%
3,000	38.07%	18.13%	5.27%
5,000	26.94%	11.72%	3.22%
10,000	15.56%	6.22%	1.64%
20,000	8.44%	3.21%	0.83%
50,000	3.55%	1.31%	0.33%
100,000	1.81%	0.66%	0.17%

Adapted from Rea and Parker (1997)

Once the appropriate sample size has been determined, the next step is to determine the individual units—households or non-household customers—that will be sampled. This brings us to a discussion of sampling frames.

6.3 Basic Sampling Frame Concepts

As defined above, a sampling frame is a list of members of a population which can be used to create a sample. If a complete and up-to-date list of all members of a population exists, this list can be used to

generate a statistically-valid sample. For PDAMs, such lists will typically be available—typically from the PDAM's finance department. In other settings, such lists do not exist and other strategies and techniques must be used to generate the sample. We will only discuss these techniques in passing.

In order to construct a sample of connected households, a PDAM might use its computerized database of connected household customers as a sampling frame. Assuming that this database is a complete listing of all connected households and does not contain a large number of households which are no longer PDAM customers, it would serve as a valid and potentially easy to use sample frame. A sample of household customers selected from this sample frame, for example, might consist of 378 households (1.89% of the PDAM's household customer base). Such a sample size would be appropriate for a population of 20,000 connected household customers with a 95% level of confidence and a confidence interval of ±5%.

To generate a sample from such a sample frame, the 20,000 households within the customer database should be assigned numbers 1 through 20,000 (or some equivalent range). 378 random numbers could then be generated and the corresponding households used as the sample. Random numbers can be easily generated using any spreadsheet's random number generator. Included within the "Customer Satisfaction Survey Implementation Manual" diskette is a spreadsheet file ("Random Sampling.xls") which contains an example of generation of a random sample from a fictional database of 500 connected household customers.

6.4 Stratified Random Sampling

This section defines stratified random sampling; discusses the motivation for stratified random sampling; provides a basic description of how stratified random sampling is carried out; and discusses some of the tradeoffs that PDAMs will face in deciding if and how to implement a stratified random sampling approach. As we did in the introduction, we remind the reader that this is not a textbook on the principles sampling and we are constrained on the amount of detailed material that we can provide. While this section adequately describes the basics of stratified random sampling, the reader should consult more specialized texts to resolve any questions that may arise which are not covered by this section.

definition and motivation

From a purely random sample selected as described in the previous section, we can make generalizations with confidence about the entire population under investigation. Frequently, however, we are also interested in understanding features of components of, or sub-populations within, the larger population. We may wish to understand the priorities of one wilayah within an urban PDAM's service area, for example. Such components of a population are termed **strata** and sampling strategies designed to ensure that they can be investigated fall under the heading **stratified random sampling**. Stratified random sampling can be an important tool for understanding subpopulations within a PDAM's service area. Implementation of stratified random sampling strategies are, unfortunately, less straightforward and normally more costly than a simple random sample of the PDAM's entire service area. If these costs and difficulties cannot be met, researchers can decide to proceed with analysis of the entire population only, or analysis of the subpopulations with a lower level of confidence.

basic principle

The basic concept underlying stratified random sampling is that, in order to make high confidence generalizations about strata *within* the overall population, we must ensure that the sample size for each stratum is sufficient. Achieving this objective can involve procedures of various cost and technical difficulty. The lowest survey-cost procedures for handling stratified random sampling unfortunately result in the most difficulties at the stage of analysis. These strategies require the use of data-weighting procedures during the analysis stage. Many statistical analysis software packages make the use of these techniques fairly straightforward. The easiest procedures to implement at the stage of analysis, on the other hand, unfortunately result in higher costs at the sampling and survey operations stage. Based on our discussions with and observations of PDAMs, we believe that in some cases it will be better to use the approach that is relatively simple to implement and results in data that are easier to analyze, but is more expensive at the survey operations phase. We will discuss this tradeoff again in general terms at the conclusion of this section.

simple example of stratification

Let's look at this procedure and the related tradeoffs for a simple example of a PDAM with 100,000 connected household customers. These customers are located in two wilayah—wilayah A has 60,000 connected households and wilayah B has 40,000 connected households. Using the basic formula for determination of required sample size, if the PDAM wishes to be 95% confident that its survey analysis results for the PDAM as a whole are within ±5% of the true value, it must sample 0.383% of its connected household customers, or 383 customers. Of these 383 customers, 60% (230) must be sampled in wilayah A and 40% (153) in wilayah B. Alternatively, we can express the ratio of wilayah A customers sampled to wilayah B customers sampled as 1.5:1.

If, however, the PDAM wishes to carry out analysis for each wilayah and be equally confident in those results, it must treat each wilayah separately. It must also ensure that the total sample contains a ratio of 1.5 wilayah A customers to 1 wilayah B customer. To determine the sample required for each wilayah, we begin with wilayah B, the smaller of the two wilayah.

wilayah B

Using the formula for sample size determination, we determine that 0.95% of wilayah B's customers must be sampled in order to produce analysis of wilayah B for which we can be 95% certain that results are within ±5% of the true value. This percentage corresponds to 380 customers. Let's look now at the wilayah A.

wilayah B

Using the formula for sample size determination, we determine that 0.64% of wilayah B's customers must be sampled in order to produce analysis of wilayah B for which we can be 95% certain that results are within ±5% of the true value. This percentage corresponds to 382 customers. Let's look now at the total combined sample for wilayah A and B.

wilayah A and B

As calculated above, we must sample 380 households in wilayah B and 382 households in wilayah A to be 95% certain, *for each wilayah considered separately*, that our results are within ±5% of the true value. What happens if we sample these numbers of households and try to draw conclusions for the entire PDAM using this sample of 762 households? We cannot do so with confidence, because we will have departed from our 1.5:1 ratio for wilayah A to wilayah B customers that we need to draw

conclusions about the population as a whole. We have many more households than are required for the PDAM as a whole, but the ratio of A:B is different from the population ratio. We face several choices.

First, we can analyze wilayah A and B separately and not consider the PDAM as a whole. This is generally not a viable option.

Second, we can use the 762 sampled households and develop a data weighting scheme which will permit us to analyze the PDAM as a whole. The use of such weighted analysis is possible, is frequently used, and is fairly straightforward using many statistical analysis software packages. It is, however, slightly more difficult analytically. Where survey cost constraints are less of an issue than analytical capacity and experience, it is in fact the preferred strategy. We recommend this approach for PDAMs that have staff with the analytical capacity to learn and use these techniques. An explanation of weighted analysis is, however, beyond the scope of this document. We refer the reader to one of the texts listed in the bibliography.

Third, we can use the required sample size for, say, wilayah B and the required A:B ratio and calculate the required sample size for wilayah A. Such a procedure works, assuming that one begins with the required sample size for the smallest stratum within the population. In our example, 380 households would be sampled in wilayah B and 570, or 1.5*380, in wilayah A. This procedure is the simplest technique which can be used that (1) maintains the required ratio between wilayah components of our overall sample, (2) ensures an adequate overall sample size, and (3) does not require the use of more complex procedures at the analysis stage. The negative for this procedure is that it can easily result in a significantly larger overall sample size.

A PDAM must consider all of these options and the relative costs and benefits of each. Overall, our recommendations are based on the financial and technical resources of the PDAM:

- PDAMs having staff with relatively high technical ability and willingness to learn software should probably use a smaller overall sample size and a weighted analysis scheme. Such an approach has both cost and capacity-building advantages.
- PDAMs having staff with lower technical ability and/or willingness to learn software should probably use a larger sample size and forego the use of weighted analysis.

6.5 Sampling in the Absence of a Proper Sampling Frame

As we have described previously, surveys are ideally based on a random or stratified random sample generated from a complete and perfectly up-to-date database of the relevant population.[1] These approaches are the sampling strategies that should be used by PDAMs. While they are the best approaches, they are sometimes not possible. What are some of the most common exceptions to sampling under this ideal setting and what strategies might be required to deal with these conditions?

[1] For this and other reasons, development of computerized databases of customers *for all PDAMs* has been a recommendation of the *WISE* project Team.

Let's look very briefly at two common cases, beginning with less problematic case and proceeding to the most difficult.[2]

case one: incomplete database on population

Common examples of incomplete databases are those where:

- data are missing for particular geographic segments of a population or
- data are missing for a particular demographic segment of a population

In the case of an incomplete set of data on the population under investigation, several strategies can be used.

First, survey designers can attempt to fill in the gaps in the available data. This approach amounts to completing the sampling frame for the areas and/or groups that are not represented. This can be an expensive and time consuming process. If survey designers are unwilling to proceed without a full sampling frame, however, this is the only option available.

A second option is to ignore the areas and/or groups that are not represented in the existing sampling frame. If such an approach is adopted, the results from the survey cannot be generalized to the population as a whole. Consider a case where complete customer data are available for 4 of the 6 kecamatans which make up a PDAM's service area. No data are available for the remaining 2 kecamatan. A survey can be carried out in the 4 kecamatan for which data area available. The resulting sample would—if properly designed and carried out—accurately represent the population of the 4 kecamatans in question, *but not the population of the PDAM*. Analysts could make statements about the 4 kecamatans, but cold not generalize to the PDAM as a whole.

A third option is to randomly sample areas for which data exist and use the procedure described immediately below for areas where data are not available.

case two: no database on population exists

The worst-case scenario is when there is no database on customers. As all PDAMs at least have paper databases of customers, this scenario typically will not be encountered. Strategies for dealing with this scenario would include the following.

- The survey team could randomly sample the entire population of that portion of the PDAM's service area for which sampling data do not exist, selecting only connected households for interviews.
- Alternatively, if the PDAM has (1) lists of customers by type *and* by some geographic breakdown of its service area—wilayah, kecamatan, kelurahan, or other subdivision—and (2) detailed maps of locations served within these areas, it is still possible to approximate a random sample. This can be done using a fairly simple, several-step process. First, the total number of customers to be interviewed must be determined. Second, street- or building-level interview assignments are made for each subdivision of the PDAM's service area. These assignments should mirror the PDAM's coverage of the area in question.

[2] Again, we remind the reader that a detailed discussion of the various sampling techniques available can be obtained from the many advanced statistics and sampling texts that are available.

7. Data Entry and Cleaning

7.1 Data Entry

Analysis of CSS data is carried out using a computer. The responses from each questionnaire must therefore be entered into a computer file of some format. (While it is possible to analyze very small amounts of data using other means, it is simply not practicable for the number of questionnaires that will constitute even a small PDAM-based CSS.)

The following procedure and accompanying example ("Data Entry.xls") are based on entry of survey data into a spreadsheet file. Spreadsheets are a commonly-used and easily understood tool that do not require extensive, specialized training to be used by most office personnel. They are not the perfect data entry platform, but their widespread use and ability to handle data easily make them an acceptable, low-cost solution.

Development and Structure of Data Entry Spreadsheet

Development of a data entry spreadsheet is the first step in the process of data entry and, in fact, typically will occur prior to administration of the survey. The spreadsheet file "Data Entry.xls" contains a simple example. The typical arrangement for a data entry spreadsheet will be a single row for each respondent and a single column, or "field," for each piece of information related to a respondent.

A data entry spreadsheet should contain the following components:

raw interview data

The cover page of all survey instruments will contain fields where enumerators enter the data which uniquely identify each respondent. These fields together form the respondent identification number. The first component of the data entry spreadsheet will be the location where data entry personnel enter this information. There should be a separate field for each component of the identification number. Entry of these fields is the first step in data entry.

constructed respondent identification number

A simple formula can be used to concatenate the individual components of the respondent identification number. (See the column labeled "RTid" in the spreadsheet file "Data Entry.xls" for an example of this concatenation formula.)

responses

Following the constructed respondent identification number, there will be fields for entering responses to each of the survey instrument's questions. There should be one field for each question.

range checks

A critical aspect of data entry and data cleaning is range checks. In a typical survey, reposes to all questions, particularly closed-ended question, will have a pre-determined maximum and minimum value. A "yes/no/don't know" question, for example, will have three possible responses. These

responses might be coded 1, 2, and 3, respectively. Responses outside of this range must be identified so that the underlying questionnaire can be checked and the entered response corrected or excluded from analysis. The file "Data Entry.xls" shows an example of simple formulae which can be used to perform range checks.

7.2 Data Cleaning

Data cleaning refers to a several-step process used to ensure that survey data have been entered correctly both by the interviewer and by data entry personnel. In the bulleted text that follows, we describe each of these simple steps in detail. We divide data cleaning into two stages, (1) data cleaning that occurs as questionnaires are returned by enumerators during the survey and (2) data cleaning that occurs after all questionnaires have been entered into a spreadsheet file.

As questionnaires are returned by enumerators, survey managers should:

- **check interview data**
  The first step in checking questionnaires as they are returned by enumerators is for the survey manager to check the interview data on the front page of each questionnaire. If the interview data are complete and legible, the questionnaire has passed the first check for completeness. The survey manager should build the full identification number on the front page of the questionnaire at this point.

  If the interview data are not complete, under some circumstances they can be reconstructed. If the enumerator only worked in one survey area—for example one kelurahan—**incomplete location data** can be easily provided. If the enumerator worked in multiple survey areas, but addresses for each interview are being recorded, incomplete location data can often be reconstructed. **Missing enumerator identification numbers** can always be entered as questionnaires are reviewed with the enumerator. If questionnaires are turned in on a daily basis (as we recommend), **missing date information** can be easily replaced. **Missing interview number data** can be easily replaced by a process of elimination. If **no interview number data** are provided for a given enumerators daily batch of spreadsheets, we can simply order the questionnaires according to the interviewers recollection of when interviews occurred. (Ordering questionnaires based on interviewer recall is less than ideal, however, because it eliminates the ability of the analysts to look for certain effects that can arise based on when the order in which interviews occur within the day. This is a very detailed type of analysis, however, and the loss would typically be quite minimal.)

- **scan questionnaires for completeness**
  In addition to reviewing interview data, survey managers should also check if questionnaires have been completed. Questionnaires that are missing responses for a substantial portion of questions should not be included in the data set. In particular, the questions that focus on customer satisfaction, customer improvement priorities, willingness to pay for improvements, and other main areas must be complete for the questionnaire to be usable for analysis. Incomplete questionnaires can either be completed on the following day by a return visit of the enumerator to the relevant interviewee or discarded.

Once the entire data file has been entered, the data cleaner—typically one of the survey analysts—should:

- **check the identification number field**
  Proper checking of questionnaires as they are returned by enumerators should catch the overwhelming majority of problems with interview data. Data entry personnel should catch the few remaining errors as they key in questionnaires. Data cleaners should, however, check identification numbers with particular care. The identification number is the piece of information that is used to group sampling units by geographic area, a critical step for analysts investigating the geographic dimensions of customer characteristics and responses.

- **perform range checks for all questions (fields)**
  All closed-ended questions and the vast majority of open-ended questions have a limited range of acceptable responses.

  Closed-ended questions are quite easy to check. When respondents answer a yes-no question, for example, their responses will be coded as "1" for "yes" and "2" for no. The data cleaner simply tabulates responses for each question or identifies the largest and smallest response for each question. The paper questionnaire for each record having an error is then pulled and the relevant question checked. The response is either corrected or, if the paper questionnaire shows the same out of range response, that respondent is excluded from analysis of that question. (Alternatively, formulae which perform range checks can be included in the data entry spreadsheet, as described above.)

  Establishing what constitutes the acceptable response range for open-ended questions is more difficult and requires the use of judgment on the part of survey data cleaners and analysts. In a survey where respondents report the number of people living in their household, for example, it would generally be logical to view a response of 233 as incorrect. What about 27, however? Or 19? Making decisions related to such borderline cases is where analyst knowledge of the particular population under investigation becomes indispensable.

- **check skip patterns**
  As described in section 3, above, the use of filter, selection, or skip questions is an important tool of the questionnaire designer. Skip patterns will not, unfortunately, always be followed correctly by enumerators. Let's think about an example. As shown in section 3, above, respondents might be instructed to skip questions related to satisfaction with PLN service if they are not a PLN customer. If a respondent states that they are not a PLN customer, but then proceeds to answer questions related to their satisfaction with PLN, data cleaners and analysts should make certain that such responses are not used in analysis. These responses would likely be the respondent's impressions of the satisfaction of PLN customers, something quite different from their own satisfaction with a service they consume.

8. Data Analysis

8.1 Overview of CSS Analysis

The current section is *by no means* an exhaustive description of all types of analysis that might possibly be done using CSS data. It is, rather, an attempt to introduce the motives for carrying out analysis and show how the process of analysis can serve these motives. We do not describe regression analysis, cross-tab analysis, factor analysis, ANOVA, or the many other tools and techniques which could be applied to CSS data. While these techniques can be used to draw out detailed and potentially interesting features of CSS data, they produce results that are less immediately useful to a PDAM. As we have already discussed, PDAM managers need analysis results which can help them act. While some might be interested in knowing, for example, the percentage of variation in satisfaction that is explained by respondent education, it will not help them act to better serve their customers.

The main way in which PDAM's will likely analyze CSS data is using simple tables and charts. Tables showing the percentage of respondents that are not satisfied with different aspects of PDAM service, for example, contain information that is immediately useful to the PDAM. Charts showing the priorities of household versus commercial customers can give the PDAM insights into the needs of different customer groups. Using such results, the PDAM can begin to get an overall view of what must be done to better meet its customers' needs. When these results are broken down by geographic area, they become even more useful—not only can the PDAM understand *what* it must do to improve service, it can understand *where* different improvements are most needed.

The following example illustrates this approach to analysis of CSS data. We begin from a very general question analyzed at the PDAM-level and proceed to more detailed questions analyzed for portions of the PDAM's service area and for specific service aspects.

8.2 An Illustration of CSS Analysis in Kotamadya Havana

Let's return to our fictional PDAM, PDAM Kotamadya Havana, for an illustration of the CSS data analysis process.

PDAM Kotamadya Havana—with exactly 100,000 household and 2,100 non-household connections—recently conducted a CSS covering all of it's six wilayah. While the survey assessed household and non-household customers and non-customers, we will focus on the connected household portion of the study in this example. The survey team wanted to be able to provide results that they were 95% confident would be within ± 5% of the true population characteristics they were attempting to measure. Based on the standard formulas for required sample sizes, the survey team calculated that they would ultimately need to interview 383 connected household customers to draw conclusions about the PDAM as a whole. Management, however, wanted to obtain results for each of the PDAMs 6 wilayah and a strategy of stratified random sampling was therefore adopted. Using the simple stratification procedure described above, the PDAM sampled 845 households in the city's 6 wilayah. A street-level sampling plan was designed using kelurahan-level data on connected customers and detailed maps of the PDAM's service areas. After one week of field operations, one

week of data entry, and several days of careful data cleaning, the team was left with a sample of 851 connected households, sufficient for the intended 95% level of confidence and ±5% confidence interval. The process carried out by the data analysts at the PDAM (using simple spreadsheets) proceeded as follows.

8.2.1 Objectives of CSS Analysis

PDAM Kotamadya Havana carried out its Customer Satisfaction Survey for several important reasons.[3]

customer satisfaction

PDAM management wanted to understand if their customers are satisfied with PDAM services. Abundant anecdotal evidence told them that many customers were not, but they required a more data-based view to guide improvement strategies. Specifically, given their fairly serious resource constraints, management wanted to be able to focus their energy and money on (1) those aspects of service which most needed improvement and (2) those portions of their service area which most needed improvement. Detailed information on customer satisfaction could help in both areas.

customer priorities

PDAM managers recognized a real need to focus their improvement efforts on service aspects which their customers felt to be most important. In the likely event that satisfaction with multiple aspects of service is low, information on priorities can help management focus on those areas which customers deem most important. As with understanding satisfaction/dissatisfaction, understanding priorities is critical to targeting resources.

willingness to pay

While some improvements in service quality might come through more efficient, more rational, or better targeted use of existing assets, many will not. Recognizing that some service improvements will likely require PDAM spending, management thought it important to develop a basic understanding of for which improvements customers are willing to pay. Such information could form the basis for more detailed, pre-investment willingness to pay surveys.

learning from other service providers and businesses

PDAM management also sought to use the CSS to identify other institutions—national utilities, local service providers, private entities, and others—from which it could learn new service practices, communications strategies, and other means of better addressing the needs of its customers.

communication

In the past, PDAM Kotamadya Havana had not made sufficient efforts to communicate with its customers. Recognizing the value that such communication can have, management sought to learn which means of communication are preferred by its customers. In addition to the benefits of better *routine* communication with customers, management had a more specific communication agenda. As with essentially all PDAMs in Indonesia, the tariff charged to household customers in Kotamadya Havana was well below a full cost-recovery level. Management wanted to learn about customer

[3] See the following section for a detailed discussion of the many possible uses of CSS data.

communication preferences so that it would be able to effectively get this message across—a message that would become particularly important in light of heightened customer expectations which would likely arise from the Customer Satisfaction Survey.

PDAM management planned to incorporate all of these pieces of information in their upcoming corporate plan and therefore wanted and *needed information on which they could act*.

8.2.2 Illustrative Analysis Process and Results

In the current section we discuss several highlights of PDAM Kotamadya Havana's analysis of its CSS and resulting action plan.

overall satisfaction results

As described above, the first information which PDAM management wanted to acquire was a basic understanding of whether its customers are satisfied with PDAM services. To begin to understand their customers' level of satisfaction, the following question was analyzed.

> **4.1.1** How satisfied are you with the quality of product and services provided by your PDAM?
> ..... _____

Analysts tabulated responses to this question for the entire PDAM, producing the following table.

PDAM Kota Havana
tingkat kepuasan pelanggan terhadap aspek pelayanan PDAM
(% dari responden mengacu kepada tingkat kepuasan)

tingkat kepuasan aspek pelayanan	memuaskan	cukup memuaskan	tidak memuaskan
Kualitas Pelayanan (UMUM)	21%	22%	57%

wilayah-level overall satisfaction results

While useful in providing an overall picture of customer satisfaction and confirming the anecdotal evidence of low satisfaction, this table reports responses for *all* household customers to a rather *general* question. This simple analysis therefore fails in terms of providing information on which the PDAM can act. To begin to move in the direction of actionable information, survey analysts constructed the following table showing overall satisfaction for each of PDAM Kota Havana's 6 wilayah.

PDAM Kota Havana
Tabel Tingkat Kepuasan Pelanggan
Kualitas Pelayanan (umum) berdasarkan Wilayah

NO.	wilayah	memuaskan	cukup memuaskan	tidak memuaskan
1	Wilayah 1	8%	44%	48%
2	Wilayah 2	6%	23%	**71%**
3	Wilayah 3	50%	26%	24%
4	Wilayah 4	31%	67%	2%
5	Wilayah 5	13%	4%	**83%**
6	Wilayah 6	5%	80%	15%

As the analysts highlighted, two wilayah stand out. In both wilayah 2 and 5, a significant proportion of respondents reported that they are not satisfied with the overall quality of PDAM product and services. (The results for wilayah 4 also stand out, with a significantly lower proportion of respondents reporting that they are not satisfied. Given PDAM Kota Havana's focus on identifying problem areas, we will ignore this result for wilayah 4.)

The analysis of KOTA Havana's CSS has begun to move toward actionable results. PDAM management now knows the two wilayah where customer satisfaction is lowest. These wilayah will be focal points of the PDAM's improvement program. But what will the PDAM focus on in these two wilayah? Why are respondents not satisfied with PDAM services? With what aspects of service are they not satisfied? To answer these important questions, we must look in more detail at the results for wilayah 2 and wilayah 5.

The questionnaire used in Kota Havana's CSS contained questions focusing not only on overall satisfaction with PDAM services, but also on more detailed aspects of service and improvement priorities. In PDAM Kota Havana's CSS questionnaire, the detailed satisfaction and improvement priority questions focused on:

- water quality,
- water quantity,
- water continuity,
- metering,
- billing procedures,
- payment procedures,
- interaction of PDAM with customers,
- communication of PDAM with customers,
- repair procedures,
- perawatan,
- sambungan baru, and

- other PDAM services.

Let's look first at the results of these more detailed questions for wilayah 2.

wilayah 2:
wilayah-level detailed satisfaction and improvement priority results

In wilayah 2, 71% of respondents reported that they are not satisfied with overall service from PDAM Kota Havana. The table below shows the results of questions covering satisfaction with particular aspects of PDAM service for wilayah 2.

PDAM Kota Havana
Satisfaction by Service Aspect
Wilayah 2

	Aspek Pelayanan:	memuaskan %	cukup memuaskan %	tidak memuaskan %
1	Kualitas Air	7%	12%	**81%**
2	Kuantitas Air	30%	17%	53%
3	Kontinuitas Air	27%	23%	50%
4	Pembacaan Meter	11%	15%	**74%**
5	Prosedur penagihan	33%	43%	24%
6	Sistem pembayaran	20%	51%	29%
7	Interaksi PDAM dengan pelanggan	39%	17%	44%
8	Komunikasi PDAM – Pelanggan	45%	40%	15%
9	Pelayanan perbaikan	10%	67%	23%
10	Perawatan	5%	80%	15%
11	Sambungan baru	44%	22%	34%
12	Pelayanan PDAM lainnya	25%	34%	41%

As this table clearly shows, connected household respondents in wilayah 2 are most likely to report dissatisfaction with two aspects of PDAM service—water quality and metering. While continuity and quantity are also a problem in this wilayah, they do not stand out like water quality and metering.

At what level of dissatisfaction for a given service aspect should a PDAM decide that action is required? This depends upon both the resources available to the PDAM and the resources required to address the problem. In an ideal world, a PDAM would have sufficient resources to focus improvements to multiple aspects of service. This situation will rarely occur, and PDAMs will be forced to make choices between competing service aspects. It is beyond the scope of this document to suggest and describe the many possible approaches for selecting between competing improvement areas. One possible approach, however, is to first focus resources on improvements in those service aspects which yield the largest increase in customer satisfaction for the smallest investment of PDAM resources. If PDAMs are subject to particular internally- or externally-determined goals, requirements, or other objectives, then it is possible that the PDAM must focus on relatively low return service aspects.

In addition to using customer satisfaction as a means of targeting resources, the management of DPAM Kota Havana utilized CSS questions and analysis related to customer priorities. The table below reports results on first improvement priority for wilayah 2.

PDAM Kota Havana
First Improvement Priority
Wilayah 2

		first improvement priority
	Aspek Pelayanan	%
1	Kualitas Air	53%
2	Kuantitas Air	2%
3	Kontinuitas Air	2%
4	Pembacaan Meter	27%
5	Prosedur penagihan	1%
6	Sistem pembayaran	4%
7	Interaksi PDAM dengan pelanggan	5%
8	Komunikasi PDAM - Pelanggan	0%
9	Pelayanan perbaikan	1%
10	Perawatan	2%
11	Sambungan baru	0%
12	Pelayanan PDAM lainnya	3%

The results of this table are not surprising. The most-frequently-occurring first improvement priority for respondents in wilayah 2 is water quality and the second most-frequently occurring first improvement priority is metering. These are the two service aspects with which respondents were least satisfied.

Let's look at the results for wilayah 5.

wilayah 5:
wilayah-level detailed satisfaction and improvement priority results

In wilayah 5, 83% of respondents reported that they are not satisfied with overall service from PDAM Kota Havana—the highest level of dissatisfaction among Kota Havana's 6 wilayah. As reported above for wilayah 2, the table below shows the results of questions covering satisfaction with particular aspects of PDAM service for wilayah 5.

PDAM Kota Havana
Satisfaction by Service Aspect
Wilayah 5

	Aspek Pelayanan	memuaskan	cukup memuaskan	tidak memuaskan
		%	%	%
1	Kualitas Air	11%	17%	**72%**
2	Kuantitas Air	7%	3%	**90%**
3	Kontinuitas Air	15%	29%	56%
4	Pembacaan Meter	23%	49%	28%
5	Prosedur penagihan	4%	72%	24%
6	Sistem pembayaran	11%	17%	**72%**
7	Interaksi PDAM dengan pelanggan	27%	35%	38%
8	Komunikasi PDAM - Pelanggan	22%	25%	53%
9	Pelayanan perbaikan	67%	13%	20%
10	Perawatan	21%	44%	35%
11	Sambungan baru	9%	56%	35%
12	Pelayanan PDAM lainnya	13%	80%	7%

Respondents in wilayah 5 report dissatisfaction primarily with three aspects of service—the quantity of water from their PDAM connections, water quality, and payment systems. Due to the PDAM's limited resources, it must focus on a smaller number of areas. For this reason, PDAM management turned to the improvement priorities in wilayah 5 as a guide. The following table reports these results.

PDAM Kota Havana
First Improvement Priority
Wilayah 5

	Aspek Pelayanan	first improvement priority
		%
1	Kualitas Air	14%
2	Kuantitas Air	**43%**
3	Kontinuitas Air	2%
4	Pembacaan Meter	5%
5	Prosedur penagihan	1%
6	Sistem pembayaran	**27%**
7	Interaksi PDAM dengan pelanggan	4%
8	Komunikasi PDAM - Pelanggan	0%
9	Pelayanan perbaikan	1%
10	Perawatan	2%
11	Sambungan baru	0%
12	Pelayanan PDAM lainnya	1%

The most frequently reported first improvement priorities in wilayah 5 are water quantity and payment systems.

comparisons to other service providers

The CSS used by PDAM Kota Havana also included questions which asked respondents to compare specific aspects of PDAM service to the same aspect of service for various other service providers. These questions were intended to identify areas where the PDAM could learn from other service providers, private enterprises, and/or other entities.

Question 6.1.3, below, is taken from the PDAM Havana questionnaire and illustrates such a comparative question.

> **6.1.3** How would you compare your satisfaction with PDAM service relative to PLN (electricity) service in terms of:
> 6.1.3.1 continuity of service ..... _____
> 6.1.3.2 billing ..... _____
> 6.1.3.3 payment systems ..... _____
> 6.1.3.4 repairs ..... _____
> 6.1.3.5 maintenance ..... _____

Based on an analysis of similar questions for PLN, PT Telkom, PN Gas, and several prominent local private enterprises, PDAM Kota Havana learned two pieces of useful information. First, a significant proportion of connected households interviewed reported that the PDAM's repair and maintenance services were significantly better than those of other (relevant) service provides. Second, many household respondents reported that PT Telkom had payment systems that are superior to those used by the PDAM.

toward an action plan for wilayahs 2 and 5

We now have a picture of (1) satisfaction by service aspect and (2) improve priorities for the two wilayah in Kota Havana where satisfaction with PDAM services is lowest. In its planning process, PDAM Kota Havana has decided to focus its limited available resources on wilayahs 2 and 5.

The PDAM's resources are limited, however, and household tariffs are low—certainly insufficient to fund improvements in the level of service. Anticipating this situation, the PDAM included questions covering household willingness to pay for improvements in different aspects of service. In the table immediately below, we report these results for both wilayahs 2 and 5.

PDAM Kota Havana
Willingness to Pay for Improvements
in Aspects of PDAM Service
Wilayahs 2 and 5

		Wilayah 2	Wilayah 5
		willing to pay	willing to pay
	Aspek Pelayanan	%	%
1	Kualitas Air	**63%**	31%
2	Kuantitas Air	31%	**74%**
3	Kontinuitas Air	11%	21%
4	Pembacaan Meter	**44%**	3%
5	Prosedur penagihan	11%	8%
6	Sistem pembayaran	4%	**35%**
7	Interaksi PDAM dengan pelanggan	0%	0%
8	Komunikasi PDAM - Pelanggan	0%	0%
9	Pelayanan perbaikan	14%	1%
10	Perawatan	12%	13%
11	Sambungan baru	5%	7%
12	Pelayanan PDAM lainnya	3%	0%

In wilayah 2, 63% of connected household respondents state that they are willing to pay for improvements in water quality, the aspect of service which they reported to be their top improvement priority. A lower, though not insignificant, number (44%) state that they are willing to pay for improvements in metering.

In wilayah 5, 74% of connected household respondents report that they are willing to pay for improvements in the quantity of water from their PDAM connection. Respondents did not report significant willingness to for improvements in any other aspect of service. (Recall that respondents in wilayah 5 also indicated significant dissatisfaction with payment systems.)

Using all of the results described above, PDAM Kota Havana developed the following action plan.

- Management at PDAM Kota Havana chose to **focus their scarce resources** on carrying out specific service improvement plans in wilayah 2 and wilayah 5. Other areas (wilayah) will be targeted in the near future, using knowledge gained from implementation of the service improvement plans in wilayahs 2 and 5.
- Given the relatively high willingness to pay for improvements in water quality (wilayah 2) and water quantity (wilayah 5), management is **undertaking a detailed study** of what system upgrades and/or modifications would be required to address these service aspects.
- Upon completion of this study, management will carry out a more detailed, **pre-project willingness to pay study** in wilayah 2 and wilayah 5.
- Management decided also to try to **create an opportunity to learn about PT Telkom's payment systems** and their implementation.

- Based on results from the communication portion of their CSS, management will make **more extensive use of radio announcements** of PDAM schedule, tariff, and repair issues.

9. Uses of Survey Data

A Customer Satisfaction Survey is a tool to help accomplish certain short-run objectives and longer-term goals. A CSS is not a goal *in-and-of-itself*. During the significant effort of designing and conducting a CSS, however, it can be easy to lose sight of the ultimate goals and objectives and become entirely focused on the survey process. It is important at this point that we step back and think about why we are doing a Customer Satisfaction Survey and what purpose it serves within the PDAM. First, let's review the important long-run goals of Customer Satisfaction Surveys which we discussed in the introduction. A CSS program can help PDAMs:

- initiate and increase *stakeholder participation* in relevant PDAM decision making processes;
- improve the *market-responsiveness* of PDAM service delivery, management, and expansion; and, ultimately,
- raise the level of PDAM *cost recovery* and of *service delivery quality*.

How can the results of Customer Satisfaction Surveys be used by PDAMs to help accomplish the important goals described above? By addressing certain short-run objectives Customer Satisfaction Surveys can help PDAMs move in the direction of addressing these longer-run goals. Let's look at these shorter-run objectives through the lens of specific uses of CSS data. We can divide the uses of Customer Satisfaction Surveys into two broad categories—internal uses and external uses. As we will see, internal uses center more on uses of data, while external uses center more on use of process. While these two views necessarily blur into each other, the division provides a useful framework for discussion.

Let's look at each in turn, citing along the way which serve each of the above long run goals, beginning with *internal* uses.

9.1 Internal Uses of CSS Results

As a component of the overall PDAM management framework, a Customer Satisfaction Survey can:

- **help identify specific areas of PDAM service that require improvement.**
  A CSS questionnaire will normally cover customer satisfaction with all areas related to the quality of PDAM service: water quality, quantity, pressure, and continuity; procedures for metering, billing, payment, complaint handling, repairs, and maintenance; contact with PDAM staff and management as well as the courtesy and professionalism of PDAM staff and management; and other relevant areas. Results related to each of these areas can be used to identify potential service aspects which the PDAM can seek to improve.

- **assist in geographic targeting of service improvements.**
  The results of a CSS can also be used to identify particular regions of a PDAM's service area that are under- or poorly-served. Respondents from one particular wilayah, while satisfied with water

quality, might desire additional payments. Respondents from kecamatans within a different wilayah might be mixed, with some satisfied with water quality and others not. When combined with information related to the PDAMs water resources, treatment facilities, and physical distribution system such data can identify physical areas where the PDAM needs to improve different aspects of its services.

- **serve as an aid in geographic targeting of service expansion.**
  While the main focus of a CSS program for most PDAMs will likely be existing customers, periodic surveys of non-customers can provide useful information as well. A survey of non-connected households, for example, can identify areas within the PDAM's service area where households wish to be connected to the PDAM system, are willing to pay connection fees, will provide labor for establishing necessary piping, and are willing to pay required tariffs. Understanding the reasons *why* households are not connected can provide the PDAM with useful information as well.

- **act as a guide for management and staff development programs.**
  A CSS should contain questions related to customer satisfaction with and impressions of PDAM staff and management. Such information can highlight potential training needs and guidelines for staff/management conduct.

- **be an important aid in the corporate planning process**
  A CSS provides PDAM management with a wealth of information related to customer satisfaction, priorities, willingness to pay for services, and preferred means of communication and interaction. This information can be a vital input into all levels of PDAM planning, from short-run activity management to longer-run strategic and corporate planning.

- **serve as a guide to investment planning**
  As with all business organizations, PDAMs routinely invest in new equipment, facilities, and other assets. These investments can take the form of upgrading existing assets on a routine basis or in an attempt to better serve the needs of their customers. Where resources are limited and customer satisfaction is a goal, the use of CSS data can help PDAMs select in which assets to invest. If only one portion of a PDAM's service area is concerned with the adequacy of payment points, for example, the PDAM can further study investing in new payment points for only that area. (*see the following item*)

- **help establish the base from which more detailed (pre-project) willingness to pay studies can be conducted.**
  A CSS can help identify service improvements and define service levels for which customers are willing to pay higher tariffs or one-time charges. A CSS as described in this document cannot, however, help PDAMs understand if customers are willing to pay a tariff that is sufficiently high to make investments in these areas cost effective. For this level of analysis, a full contingent valuation or willingness to pay study is required. A CSS can serve, however, serve as an important step in specifying the service level or improvements on which such a study is based. Recent literature has, in fact, called into question the success of many pre-project willingness to pay studies precisely because of insufficient stakeholder involvement in specification of the service level analyzed.

- **be used by PDAM managers to help push staff in the direction of thinking and acting in a more customer-oriented manner.**
  To design and conduct a Customer Satisfaction Survey is, by definition: to *think* about customers, to *interact* with customers, to *obtain information* from customers, and to begin to *use information and analysis* to address customer concerns in a "better" and/or more appropriate manner. If properly used within the PDAM, all of these activities can help all staff develop a more customer-focused approach. PDAM management should strive to involve as many managers and staff members as possible in defining the objectives of a CSS. PDAM managers should routinely discuss the results of the CSS in staff meetings and management strategy session.

- **can identify other service providers and/or businesses with whom PDAMs can develop information-sharing relationships.**
  A common feature of Customer Satisfaction Survey questionnaires is questions which ask respondents to compare the performance of one service provider to another across various dimensions. Do PDAMs perform as well as electricity providers in terms of billing and payment procedures? … as well as telephone companies in terms of customer complaint handling? … as well as private businesses in terms of developing new products, services, or procedures to meet their evolving needs? By learning which entities do a better job in which dimensions of service, PDAMs can identify sources of information about new or improved business practices which can help them better serve their customers.

In addition to these internal benefits, a Customer Satisfaction Survey program can be a valuable aid in building the PDAM's relationships with its customer base and other stakeholders. We discuss these external benefits

9.2 External Uses of CSS Results

As described above, a Customer Satisfaction Survey program can serve valuable internal objectives—objectives that center on PDAM focus, learning, and behavior. A CSS can also serve more externally-focused ends as well. Specifically, a CSS program can:

- **help establish relationships with consumer groups.**
  The process of conducting a CSS will necessarily involve significant interaction with *individual* consumers. All stages of a CSS can be designed to include the participation of organized consumer *groups* as well: pre-survey focus groups can be organized by local groups, members of consumer groups can be actively recruited and used as enumerators, consumer groups can help establish the study's analysis objectives, and results of the CSS can be disseminated through consumer groups. All of these actions can help PDAMs build and maintain important relationships.

- **provide customers with an indirect vehicle through which they participate in PDAM management and begin to establish the relationships with customers and customer groups that can lead to more direct participation in PDAM management.**
  A CSS is a means of transferring information from customers to the PDAM that serves them. By providing information on satisfaction, priorities, and other areas covered by a CSS, customers can

begin to influence the actions and directions taken by the PDAM. While not a direct/personal input into management discussions and decisions, indirect input can be a first step in demonstrating the value *to the PDAM* of customer voice. Establishing the value of customer input can help build PDAM willingness for a more direct and routine customer role in management.

- **help identify customers' preferred forms of communication with PDAM management and staff.**
  A key component of any Customer Satisfaction Survey will be investigation of customers' preferred means of communicating and interacting with the PDAM. It is perhaps a hackneyed saying, but effective communication is a two-way street—providing information *to* customers and learning from consumers are both critical functions which support PDAM planning and routine behavior. Understanding how customers want this interaction to occur will make customers more willing to communicate, will make PDAMs better able to reach their customers, and will contribute to the PDAM-customer relationship.

Bibliography and Sources for Additional Reading

Deaton, Angus (1997), The Analysis of Household Surveys: A Microeconometric Approach to Development Policy, Johns Hopkins University Press, Baltimore.
A detailed presentation and explanation of advanced analysis survey analysis techniques, focusing on investigating the micro-economic behavior of households.

Hayes, Robert E. (date), Measuring Customer Satisfaction: survey design, use, and statistical analysis methods, publisher, location.
Covers basic customer satisfaction concepts, but contains only a weak treatment of survey design, sampling, and analysis topics.

Magnani, Robert (1999), Sampling Guide, Food and Nutrition Technical Assistance Project (*FANTa*), United States Agency for International Development.
A good, brief treatment of more sophisticated sampling concepts. While focused on food and nutrition applications, basic and more-sophisticated sampling concepts are presented in a compact, readable format.
HTML version available at http://www.fantaproject.org/docs/sampling.htm

Rea, Luis M. and Robert A. Parker (1997), Designing and Conducting Survey Research: a comprehensive guide, Jossey-Bass Publishers, San Francisco.
A good general source on survey research, covering topics ranging from the principles of questionnaire design to basic survey analysis. While not a rigorous statistical/mathematical text, the important principles of survey-related statistics and associated formulae are presented, discussed, and used in illustrative examples.

Appendix One
Materials for Informal Customer Focus Groups

Appendix Two
Enumerator Training Outline

Appendix Three
Enumerator Training Outlines with Notes

Appendix Four
Survey Field Log

Appendix Five
Pilot Survey Questionnaires