

An Overview to Survey Tools and Techniques for Quantitative Research

Quantitative Research techniques involve relatively large number of respondents and are designed to generate information that can be projected to the whole population using a representative sample.

I. Survey

A. Survey: Context

- It is a structured self-completion questionnaire given to a sample of a population and designed to elicit specific information from the respondents.
- Often demographic and lifestyle characteristics form part of the surveys and may be compared to respondents' attitudes, intentions, etc.
- The survey technique can be used for deductive analysis and inductive analysis. In the former the research team can start with a hypothesis, theory or model and then collect data to test it. In the latter, the research team can start collecting data and then infer a hypothesis, theory or model from patterns that emerge from the data analysis.
- The research team must identify the kind of information it needs to gain solutions or insights on the problem/project and achieve its research/project objectives, and retain that focus throughout the survey design process.
- The research team needs to think ahead about how it would collate and analyse the information gathered using the survey at the stage when the survey is being designed. Designing a survey that produces a range of information that is difficult to collate in any meaningful quantitative or qualitative way would defeat the purpose of the data collection exercise.

B: Survey: Content

- No matter how good the survey design and how rigorous the data analysis, a poorly worded survey will negate these efforts.
- The validity (the extent to which the data accurately measures what they were intended to measure) and reliability (the extent to which the data collection method will yield consistent findings if replicated by others) of

the data the research team collects depends on the design of the survey and the choice of words and phrases.

- A survey should be very respondent friendly. The order and flow of questions should be logical to the respondent. A sequence of questions that moves from the general to the specific would be desirable.
- A blend of open-ended and close-ended questions is desirable. Too many open-ended questions would be a time-consuming exercise for the respondent. Too few open-ended questions would leave them with a sense of inability to contribute enough to the survey and the study.

C. Survey: Pilot Testing

- All surveys should be piloted with a small group before the main research to assess their value, validity, and reliability.
- The pilot testing of the survey instrument should be done under conditions as similar with the upcoming fieldwork.
- Use of respondents like those in the planned fieldwork would be desirable for the pilot testing exercise.
- The feedback received during pilot testing will enable the research team to know whether the survey fulfils the objectives of the data collection exercise. Some of the points that the research team should focus on during the pilot testing are:
 - How easy was it to follow the survey instructions?
 - How effective was the layout and format?
 - Were all the questions free from ambiguity?
 - Were any questions too sensitive or embarrassing for the respondent?
 - Where there any questions that the respondent found to be irrelevant?
 - Were there questions left without a response?
 - Were there too many questions with 'Don't know' as the answer?
- Answers to the above-mentioned questions during the pilot testing phase would help the research team redesign/modify the survey instrument.

D. Survey: Response

- There can be a low rate of return with surveys. Hence, the research team needs to introduce the study and the survey carefully and courteously to potential respondents. This introduction can include the use of a covering letter that provides the context of the study/project, the relevance of the

respondent's experience to the area of study/project, and the tentative time it will take to respond to the survey.

- The research team should also specify to the respondent how and in what format it will disseminate the information gained through the survey.
- Often, offering an inducement can also improve the rate of return of questionnaires. There is a scope for lateral thinking on what would make potential respondents inclined to participate in the survey.

For e.g., sharing the survey findings/synopsis of study with the respondents would also be a good gesture by the research team for the time invested by the respondents in replying to the survey, which has benefited the study/project. A gift/monetary contribution/donation to the respondent's chosen cause would be another possibility.

E. Survey: Uses

- Surveys can be used to measure:
 - Attitudes (likes and dislikes: 'Are you happy that ____?')
 - Preferences ('Do you prefer A or B?')
 - Beliefs ('Do you believe that ____?')
 - Behavioural Experiences ('How often do people you know ____?')
 - Descriptions ('How many times have your employees ____?')
- Surveys can be used to find:
 - Prevalence of characteristics, attitudes, preferences, beliefs, behaviour (cross-sectional)
 - Changes in these over time (longitudinal)
 - Differences between groups of people based on geography, demography, sociology, etc.

II. Variables

These help in measuring the responses received to each of the questions/ each category of questions used in the survey. Variables permit conversion of propositions to hypotheses and enable testing them.

Types of Variables

A. Nominal Variables

- A variable with values which have no numerical value, such as gender or occupation.
- Labels having no quantitative value or origin.
- Mutually exclusive + exhaustive set of categories
E.g.: Distinct categories such as age groups, religious affiliation, region of origin, etc.

B. Ordinal Variables

- A variable with values whose order is significant, but on which no meaningful arithmetic-like operations can be performed.
- As nominal + indicator of order
E.g.: Where preferences, etc. can be ranked (Fast < Very Fast < Very, very fast)

C. Interval Variables

- An ordinal variable with the additional property that the magnitudes of differences between two values are meaningful.
- Power of nominal + ordinal + concept of equality of interval.
Ex 1: Intrinsically numeric but lacking a real 'zero' point; e.g. 'temperature,' where 40 degrees is not twice as warm as 20 degrees
Ex 2: Absolute time:
10PM (today) > 8PM (today)
10PM (today) - 8PM (today) = 2 hours
8PM (today) + 2 hours = 10pm (today)
2PM + 10PM - not defined
10PM / 8PM - not defined
10PM * 8PM - not defined
10PM * 2 hours - not defined

D. Ratio Variables

- A variable with the features of interval variable and, additionally, whose any two values have meaningful ratio, making the operations of multiplication and division meaningful.
- Represents actual amount of a variable. Permits division/multiplication.
- Power of all + origin.

Ex 1: Intrinsically numeric and having a real 'zero' point; e.g., length, where 40 cm is twice as long as 20 cm

Ex 2: Speed

10 meters per second > 8 meters per second

$10 \text{ mps} - 8 \text{ mps} = 2 \text{ mps}$

$8 \text{ mps} + 2 \text{ mps} = 10 \text{ mps}$

$10 \text{ mps} / 8 \text{ mps} = 1.25$ (times as fast)

$1.25 * 8 \text{ mps} = 10 \text{ mps}$

III. Survey: Question Types

There are a variety of ways in which survey questions can be designed. The choice of type of questions depends on the research objectives and the insights for which the data is being gathered. A list of the various types of questions that can be used while designing a survey has been provided here using the example of Air India and the civil aviation industry. The kind of survey variables to be used for each type of question has been listed in brackets.

- **Dichotomous (Nominal)**

In arranging this trip, did you contact Air India?

Yes No

- **Multiple Choice - Single Response (Nominal)**

With whom are you traveling on this trip?

- No one
- Spouse
- Spouse and children
- Children only
- Business associates/friends/relatives
- An organised tour group

- **Multiple Choice - Multiple Response (Nominal)**

Which airlines do you usually travel by?

- Air India
- Vistara
- Spice Jet
- Indigo
- Go Air

- **Likert Scale (Interval)**

Indicate your level of agreement with the following statement:

‘Small airlines generally give better service than large ones.’

- Strongly Disagree (SD)
- Disagree (D)
- Neither Agree/Nor Disagree (NA/ND)
- Agree (A)

Strongly Agree (SA)

▪ **Semantic Differential Scale (Interval)**

Air India

LargeSmall

Experienced.....Inexperienced

Modern.....Old-fashioned

▪ **Importance Scale (Interval)**

Airline food service is _____ to me.

- Extremely important
- Very important
- Somewhat important
- Not very important
- Not at all important

▪ **Rating Scale (Interval)**

Air India's food service is _____.

- Excellent
- Very good
- Good
- Fair
- Poor

▪ **Numerical Scale (Ordinal/Interval)**

Extremely Good 5 4 3 2 1 Extremely Bad

Air India

Employee courteousness at the time of check in: __

Employee's cooperation on board the flight: __

Employee' knowledge of task: __

▪ **Intention to Buy Scale (Nominal)**

How likely are you to purchase tickets of Air India if in-flight Internet access were available?

- Definitely buy
- Probably buy
- Not sure

- Probably not buy
- Definitely not buy

- **Multiple Rating Scale (Interval)**

Please indicate how important or unimportant each of the following is:

	Unimportant 1	2	3	4	5	6	Important 7
Ticketing Convenience							
Seating							
Quality of Food							

- **Fixed Sum (Ratio)**

Taking all the service characteristics what is their relative importance to you?
(Divide 100 units between):

Being one of the lowest cost airlines: ___ points

All other aspects of airline performance: ___ points
100 points

- **Completely Unstructured**

What is your opinion of Air India?

IV. Survey: Question Phraseology

The choice of words and phrases used in designing the survey plays an especially important role in the type of data gathered using the survey instrument. Six types of faults are often found in survey questions. These include:

- **Bias:** Likely to lead to distortion or prejudice
- **Ambiguity:** Having more than one meaning
- **Jargon:** Use of words that may not be widely understood
- **Knowledge:** Unwise assumption about respondents' knowledge
- **Insensitivity:** Showing lack of concern for respondents' feelings
- **Leading:** Prompts or encourages the respondent with the bias of the research team

Examples of Imperfect Questions and Suggested Solutions

Imperfect Question	Fault and Suggested Solution
i. Are you against giving too much power to trade unions?	i. <u>Bias</u> : Amend to: 'What are your views on the power that trade unions have?' Can offer respondent several alternative responses
ii. What is your social class? Upper/Middle/Lower	ii. <u>Bias</u> : Amend to: 'What is your employment category?' Offer respondent several alternative responses in alphabetic order
iii. 'Stakeholder Capitalism can help achieve a better bottom line, better employee loyalty, and better customer satisfaction.'	iii. <u>Ambiguity</u> : Divide the question into three separate questions about the relationship between Stakeholder Capitalism and better bottom line, better employee loyalty and better customer satisfaction. Can use a five-point Likert Scale from 'strongly disagree' to 'strongly agree' to gain answers for each question.
iv. What is your job?	iv. <u>Ambiguity</u> :

	<p>Make clear what sort of data the research team wants about the person's job.</p> <p>For e.g., job title, type of job - technical/managerial, job responsibilities</p>
<p>v. 'Over the past three years, the financial performance of my company in comparison with the average in my industry was:' (Use of a 7-point scale from 'significantly worse' to 'significantly better' for the response)</p>	<p><u>v. Ambiguity:</u></p> <p>'Financial performance' is not well-defined.</p> <p>Amend to: 'Over the past three years, the percentage increase in market capitalization...'</p> <p>(In the response options, 'Significantly' is very subjective!</p> <p>Amend response options to: 'More than 50% above average, 25-50% above average, 1-25% above average, 0-25% below average....')</p>
<p>vi. 'I am not the kind of leader who would not make organisational culture a priority.'</p>	<p><u>vi. Ambiguity:</u></p> <p>Avoid use of double negatives</p> <p>Amend to: 'I am a leader who places considerable importance on organisational culture.'</p> <p>(Can provide a five-point Likert Scale from 'strongly disagree' to 'strongly agree' for response options.)</p>
<p>vii. Does your organisation align with the ESG priorities highlighted in the United Nations Sustainable Development Goals?</p>	<p><u>vii. Jargon:</u></p> <p>ESG is a jargon and should be explained.</p> <p>Amend to: 'ESG priorities identified by the United Nations Sustainable Development Goals underscore the importance of focusing on Environmental, Social and Governance goals during business decision making.... Does your organisation align with these?'</p>

	Can specify a range of responses, including 'Don't know'
viii. Do you agree with your company's mission statement?	viii. <u>Knowledge assumed</u> : It is assumed that the respondent is familiar with her/his company's mission statement. Make two questions: 'Are you familiar with your company's mission statement?' 'If so, do you approve it?'
ix. Man Management is key for every manager – Agree?	ix. <u>Insensitive</u> : Amend to: 'Managing people is the key to success for every manager. Do you agree?'
x. Are you: Married/Separated/Divorced?	x. <u>Insensitive</u> : Amend to: 'Please indicate your status.' Specify appropriate categories for response.
xi. Do you read a management journal?	xi. <u>Leading</u> : Amend to: 'Indicate if you read management journals regularly' Specify several journals for responses.
xii. 'Progressive executives appreciate the importance of spirituality in business leadership.'	xii. <u>Leading</u> : Amend to: 'I appreciate the importance of spirituality in business leadership.' Can provide a five-point Likert Scale from 'strongly disagree' to 'strongly agree' for response options.

V. Sampling Techniques for Primary Data Collection through Surveys

A. Key Terms

- A *Population* is the total collection of elements about which we wish to make some inferences.
For e.g., all office workers in a firm, or all students in a University, compose a population of interest.
- A *Population Element* is the subject on which measurement is being taken. It is the unit of study.
For e.g., each office worker questioned about a flexible work schedule, or each student questioned about infrastructure facilities on the university campus, is a population element.
- A *Census* is a count of all the elements in a population.
For e.g., Data collected from every Indian citizen during population census conducted once every decade.
- A *Sample* is a selected subset of the Population. The basic idea of Sampling is that by selecting some of the elements in a population, the research team can draw conclusions about the entire population.

B. Characteristics of a Good Sample

A sample should be selected based on the following criteria:

- **Accuracy**
 - Accuracy is the degree to which bias is absent from the sample.
 - When the sample is drawn properly, some sample elements underestimate the population values being studied and others overestimate them. Variations in these values offset each other; this counteraction results in a sample value that is generally close to the population value. For these off settings to occur, there must be enough elements in the sample, and they must be drawn in a way to favour neither overestimation nor underestimation.
For e.g., If interested in value of homes in South Mumbai, do not sample homes with a better view for e.g., sea facing flats in South Mumbai, that tend to be more valuable

- **Precision**

Precision of estimate is the second criterion of a good sample. No sample will fully represent its population in all respects. The numerical descriptors that describe samples may be expected to differ from those that describe populations because of random fluctuations inherent in the sampling process. This is typically measured by Standard Error of Estimate (Type of Standard Deviation Measurement)

C. Steps in Sampling Design

1. Relevant Population

The definition of the population may be apparent from the problem/project or the research question(s).

2. Parameters of Interest

Population parameters are summary descriptors of variables of interest in the population.

For example

- i. Attendance at a special event – Participation in a promotion (yes/no)
- ii. Percent of employees who visit the gym in the office – Frequency of visit (very often, medium, rarely)
- iii. Mean temperature of ideal vacation location – Temperature in degrees
- iv. Average number of store visits per month – Actual number of store visits

3. Sampling Frame

The sampling frame is closely related to the population. It is a complete and correct list of all elements from which the sample is drawn.

4. Type of Sample

This refers to the choice made by the researcher – whether to use a probability or non-probability sampling.

5. Sample Size

Principles that influence sample size include:

- i. The greater the dispersion or variance within the population, the larger the sample must be to provide estimation precision.

- ii. The greater the desired precision of the estimate, the larger the sample must be.
- iii. The narrower the interval range, the larger the sample must be.
- iv. The higher confidence level in the estimate, the larger the sample must be.
- v. The greater the number of subgroups of interest within a sample, the greater the sample size must be, as each subgroup must meet minimum sample size requirements.
- vi. If the calculated sample size exceeds 5% of the population, sample size may be reduced without sacrificing precision.

6. Cost

Cost considerations influence decisions about the size and type of sample and also the data collection methods. Almost all studies have some budgetary constraint, and this may encourage a researcher to use a non-probability sample. Probability sample surveys incur a variety of other costs that are not necessary when arbitrary methods are used.

D. Sampling Techniques

There are two main types of sampling techniques:

1. Non-Probability Sampling

- It is a technique that does not use chance selection procedures to identify those who will participate. It relies on the personal judgment of the research team to decide who will be included in the sample.
- It is used when the research team has little initial control over the choice of who is presented for selection, or where controlled selection of participants is not a critical factor.

Non-Probability Sampling Techniques

i. Convenience Sampling:

Sampling those most convenient; those immediately available.

E.g., survey responses from friends about the user friendliness of features of a social media platform

ii. Voluntary Sampling:

The sample is self-selecting. Respondents come forward voluntarily in response to an appeal.

E.g., people replying to an ad that promised them benefits if they reply to the survey.

iii. Purposive / Judgmental Sampling:

It enables the research team to use its judgment to choose people that are presented or are available that best meet the research team's objectives.

For e.g., Researchers may seek respondents whom they feel are part of the target population.

iv. Quota Sampling:

This is a restricted judgmental sampling technique in which the first stage consists of developing control categories or quotas of the population and the second stage consists of sample elements being selected based on convenience or judgment.

For e.g., Deciding that there will be a quota of 20 respondents aged in their 20s, 20 aged in their 30s and so on.

v. Event Sampling:

It involves using the opportunity presented by a particular event.

For e.g., A religious/cultural festival like Kumbh Mela to study public hygiene and sanitation infrastructure.

vi. Snowball Sampling:

It is a sampling technique in which an initial group of respondents is selected randomly, and then subsequent respondents are identified based on the referrals provided by the initial respondents. It is used when respondents are difficult to identify and can best be located through referral networks.

For e.g., A sample of respondents is obtained, and then subsequent respondents are accessed based on the referrals/network of the first set of respondents.

2. Probability Sampling

- It is a sampling technique in which each element of the population has the same probabilistic chance of being selected for the sample.

- It is used when the research team has a significant measure of control over who is selected for responding to the survey and on the selection methods for choosing the respondents.

Probability Sampling Techniques

i. Simple Random Sampling:

It is a technique in which each member in the population has a known and equal probability of selection. Every element is selected independently, and the sample is drawn by random procedure from a sampling frame.

For e.g.,

Population: 5,000 full-time students at a university.

Frame: University records

Sample Size: 100

Simple Random Sampling: List all in frame, assign numbers from 1-5000, select 100 students using random numbers

ii. Systematic Sampling:

It is a technique in which the sample is chosen by selecting a random starting point and then picking every n th element in succession from the sampling frame.

For e.g., In the earlier example, choosing the first name on the list using a random number and then selecting every 50th name thereafter.

iii. Stratified Sampling:

- It is a two-step process that partitions the population into sub-populations or strata after which elements are selected from each stratum by a random procedure.
- There is greater homogeneity within each stratum and heterogeneity between them.

For e.g., In a study of an organisation that has 50 supervisors and 800 labourers, a 10% representative sample of this population using stratified sampling technique would target 5 supervisors and 80 labourers for the survey.

iv. Cluster Sampling:

- It is another two-step technique by which the target population is divided into mutually exclusive and collectively exhaustive subpopulations called clusters, and then a random sample of clusters is selected based on a probability sampling technique such as simple random sampling.
- There is greater heterogeneity within subgroups, and greater homogeneity between them.

For e.g., An interview survey of about 120 households in a small town is to be taken. Using Cluster Sampling, the town could be divided using a map, into 100 clusters of about 30 households each. Then the research team can randomly select 4 of the clusters and survey all the households in those 4 clusters.

E. Sampling Errors

- An important part of data collection is trying to minimize the incidence and effect of errors on research.
- The total error is composed of two main parts:
 - i. Random Sampling Error
 - ii. Non-Sampling Error

1. Random Sampling Error

- This occurs when the sample members selected for the study do not represent the population of interest. Therefore, the responses are not coming from those whom the team really wants to research.
- This can be due to carelessness on the part of the research team members as it is up to them to select the correct population and truly representative sample.

2. Non-Sampling Error

Non sampling errors do not relate to the way the research team selects the respondents. These are of two types:

i. Response Errors

Response Errors occur when the respondents give accurate answers, or their answers are mis-recorded or misanalysed and can be caused by the research team or the respondents. These are of two types:

- **Researchers Errors** can include:
 - a. **Surrogate Information Error:** Here the research team asks a wrong question and thus does not get the information needed to answer the research question. E.g., Asking about the consumer preferences when what is really needed is purchase intention.
 - b. **Measurement Error:** Measuring the information in a wrong way. E.g., Using a scale to measure perceptions rather than preferences.
 - c. **Population Definition Error:** This is the incorrect interpretation of what population has to be researched
 - d. **Sampling Frame Error:** This is the difference between the population defined and the sampling frame actually used.
 - e. **Data Analysis Error:** This is the incorrect interpretation of the research findings.

- **Response Error** can include:
 - a. **Inability Error:** This occurs when the respondent is unable to provide accurate answers to the question(s)
 - b. **Unwillingness Error:** This occurs when the respondent is unwilling to provide accurate information. E.g., trying to give socially acceptable answers to avoid embarrassment.
 - c. **Bias:** This is caused by factors that the research team cannot control resulting in the respondents changing their answer(s). For e.g., male respondents may give different answers depending on whether the interviewer is a male or a female.

ii. Non-response Errors

- This occurs when respondents included in the sample do not respond to the questionnaire or a particular question. This is beyond the research team's control and is usually due to either refusal to participate or the respondents not being available.
- If this situation occurs, the research team must find some willing respondents in the same sample population, change the sample population, change the wording of any ambiguous questions, or at worst, cancel the research design and discover another way of obtaining the same data.