# Data and data Collection Methods

Lecturer: Ralston Henry

# Objectives

❖ Differentiate Between Primary and Secondary Data.

❖ Asses the Advantages and Disadvantages of Primary and Secondary Data.

❖ Outline the different Data collection Methods.

❖ Outline the Strength and Weaknesses of various data collection tools

# Data Collection Methods

## Pros and Cons of Primary and Secondary Data

# Where do data come from?

❖Typically data can be categorized in two groups based on the origin or source of the data.

➢Secondary data

➢Primary data

# Secondary Data

❖ Secondary data – This is data that someone else has collected.

❖ This is typically found in:

➢ Vital Statistics – birth, death certificates

➢ Hospital, clinic, school nurse records

➢ Private and foundation databases

➢ City and county governments

➢ Surveillance data from state government programs

➢ Federal agency statistics - Census, NIH, etc.

# Secondary Data – Limitations

❖ What did you find frustrating when looking for data on the websites?

# Secondary Data – Limitations

❖ When was it collected?  For how long?

➢ May be out of date for what you want to analyze.

➢ May not have been collected long enough for detecting trends.

▪ E.g. Have new anticorruption laws impacted Russia's government accountability ratings?

# Secondary Data – Limitations

❖ Is the data set complete?

➢ There may be missing information on some observations

➢ Unless such missing information is caught and corrected for, analysis will be biased.

# Secondary Data – Limitations

❖Are there confounding problems?

➢Sample selection bias?

➢Source choice bias?

➢In time series, did some observations drop out over time?

# Secondary Data – Limitations

❖Are the data consistent/reliable?

➢Did variables drop out over time?

➢Did variables change in definition over time?
  • E.g. number of years of education versus highest degree obtained.

# Secondary Data – Limitations

❖ Is the information exactly what you need?

➢ In some cases, may have to use "proxy variables" – variables that may approximate something you really wanted to measure. Are they reliable? Is there correlation to what you actually want to measure?

➢ E.g. gauging student interest in U.W. by their ranking on FAFSA – subject to gamesmanship.

# Secondary Data – Advantages

❖No need to reinvent the wheel.

➢If someone has already found the data, take advantage of it.

❖It will save you money.

➢Even if you have to pay for access, often it is cheaper in terms of money than collecting your own data. (more on this later.)

# Secondary Data – Advantages

❖ It will save you time.

➢ Primary data collection is very time consuming.

❖ It may be very accurate.

➢ When especially a government agency has collected the data, incredible amounts of time and money went into it. It's probably highly accurate.

❖ It has great exploratory value

➢ Exploring research questions and formulating hypothesis to test.

# Primary Data

❖ Primary data – This is data that you collect.

❖ This is typically acquired via:

➢ Surveys

➢ Focus groups

➢ Questionnaires

➢ Personal interviews

➢ Experiments and observational study

# Primary Data - Limitations

❖ Time is not always available to:

  ➢ Designing collection instrument?

  ➢ Selecting population or sample?

  ➢ Pretesting/piloting the instrument to work out sources of bias?

  ➢ Administration of the instrument?

  ➢ Entry/collation of data?

# Primary Data - Limitations

❖ **Uniqueness**

➢ May not be able to compare to other populations or findings in other from other researches.

❖ **Possibility of Researcher error**

➢ Sample bias

➢ Researcher subjectivity

➢ Other confounding factor

# Data collection choice

❖ **What you must ask yourself:**

➤ Will the data answer my research question?

❖ **To answer that question the researcher must:**

➤ First decide what your research question is and what data/variables are needed to scientifically answer the question

# Data collection choice

❖If that data exist in secondary form, then use them to the extent you can, keeping in mind the limitations of the data and source.

❖But if it does not, and you are able to fund primary collection, then it is the method of choice.

# Data Collection

What is the best way to go about collecting data?

O'Leary, Z. (2004) The Essential Guide to Doing Research. London: Sage

fppt.com

# The Quest for Credible Data

❖Virtually all methodologies are reliant on the collection of credible data

❖The first step in any form of data collection is gaining legitimate access to the source of the data

fppt.com

# Surveys

❖ ***Survey (n):*** Information gathered by asking a range of individuals the same questions related to their characteristics, attributes, how they live, or their opinions

❖ ***Survey (v):*** The process of collecting such information

# Knowing what to expect

❖A good survey has the potential to reach a large number of respondents; generate standardized, quantifiable, empirical data - as well as some qualitative data; and offers confidentiality / anonymity


❖Credible data, however, can be difficult to generate

fppt.com

# Basic Survey Types

❖ Surveys can be:

➢ ***Descriptive:*** These surveys pretty much do what they say - they describe. The goal is to get a snapshot - of your 'respondents'

➢ ***Explanatory:*** These surveys go beyond description (although they do gather descriptive data) and attempt to establish why things might be the way they are

fppt.com

# Basic Survey Types

❖ Surveys can also involve populations or samples of populations:

  ➢ *Census:* This is a survey that does not rely on a sample. A census surveys every single person in a defined or target population

  ➢ *Cross-sectional surveys:* This type of survey uses a sample or cross-section of respondents selected to represent a target population

24

# Basic Survey Types

❖ Surveys can capture a moment or map trends:

> ➤ **_Trend surveys:_** A trend survey asks similar groups of respondents, or the same cross-section, the same questions at two or more points in time

> ➤ **_Panel study:_** A panel study involves asking the same (not similar) sample of respondents the same questions at two or more points in time

fppt.com

# Basic Survey Types

❖ Surveys can be administered:

➢ Face to face

➢ Telephone

➢ Self-administered

# Conducting a Survey

❖Conducting a survey capable of generating credible data requires:

➢thorough planning

➢meticulous instrument construction

➢comprehensive piloting

➢reflexive redevelopment

➢deliberate execution

➢and appropriate analysis

# Questions to Avoid

- ➢ Poorly worded questions:
    - Complex terms and language
    - Ambiguous questions
    - Double-barreled questions
- ➢ Biased/leading or loaded:
    - 'Sound true' statements
    - Hard to disagree with statements
    - Leading questions
- ➢ Problematic for the respondent:
    - Recall dependent questions
    - Offensive questions
    - Questions with assumed knowledge
    - Question with socially desirable responses

# Response Categories

❖ Survey questions can either be open or closed:

➢ ***Open questions:*** These questions ask respondents to construct answers using their own words. Open questions can generate rich and candid data, but it can be data that is difficult to code and analyse

➢ ***Closed questions:*** These questions force respondents to choose from a range of predetermined responses, and are generally easy to code and statistically analyse

# Response Categories for Closed Questions

➤ **Yes / No - Agree / Disagree**:

   *Do you drink alcohol?*                                    *Yes/ No*

➤ **Fill in the blank:**

   *How much do you weigh?*                    _____

➤ **Choosing from a list:**

   *What would you drink most often?*

➤          *Beer            Wine                Spirits            Mixed drinks        Cocktails*

➤ **Ordering options:**

   *Please place the following drinks in order of preference*

➤          *Beer            Wine                Spirits            Mixed drinks        Cocktails*

➤ **scaling:**

   *It is normal for teenagers to binge drink*          *1        2        3        4        5*

                                           *strongly disagree   disagree    unsure    agree    strongly agree*

# Additional Considerations in Survey Construction

❖ Providing clear background information and lucid instructions

❖ Logical organization

❖ Comprehensive coverage without undue length

❖ User friendly and aesthetically pleasing layout and design

# Interviewing

❖ ***Interviewing:*** A method of data collection that involves researchers asking respondents basically *open ended* questions of respondents.

❖ Interviews are capable of generating both standardized quantifiable data, and more in-depth qualitative data.

❖ However, the complexities of people and the complexities of communication can create many opportunities for miscommunication and misinterpretation.

# Interview Types

❖Interviews can range from fixed to free:

➤***Structured:*** Uses pre-established questions, asked in a predetermined order, using a standard mode of delivery

➤***Semi-structured:*** As the name suggests, these interviews are neither fully fixed nor fully free, and are perhaps best seen as flexible

➤***Unstructured:*** Attempts to draw out information, attitudes, opinions, and beliefs around particular themes, ideas, and issues without the aid of predetermined questions

# Interview Types

❖ Interviews also vary in their level of formality:

➢ *Formal:* A formal interview is just that, formal. Perhaps the best analogy is the classic job interview that includes: the office setting; the formal handshake; appropriate attire; order and structure; and best professional behaviour.

➢ *Informal:* An informal interview attempts to ignore the rules and roles associated with interviewing in an attempt to establish rapport, gain trust, and create a more natural environment conducive to open and honest communication.

# Interview Types

❖ Interviewing can be done one on one, or in groups:

➢ ***One on one:*** Most interviews are an interaction between the interviewer and a single interviewee. It is thought that 'one on one' allows the researcher control over the process and the interviewee the freedom to express their thoughts

➢ ***Group:*** Group interviews involve interviewing more than one person at a time. This can be done in a formal structured way, or may involve a less structured process where the researcher acts more as a moderator or facilitator than an interviewer

# Conducting an Interview

❖ Conducting an interview that can generate relevant and credible data requires:

- ➤ thorough planning
- ➤ considered preparation of an interview schedule and recording system
- ➤ sufficient piloting
- ➤ reflexive modification
- ➤ and appropriate analysis

fppt.com

# Conducting the Interview

❖Two important considerations in interviewing are:

➢ Presentation of self

➢ Preliminaries

- Be on time!
- Set up and check equipment
- Establish rapport
- Introduce the study
- Explain ethics

# The Questioning Process

❖ **Listen more than talk**

➢      The main game in interviewing is to facilitate an interviewee's ability to answer. This involves:

- easing respondents into the interview
- asking strategic questions
- prompting and probing appropriately
- keeping it moving
- being true to your role
- winding it down when the time is right

# Recording Responses

❖ Recording responses can be done in a number of ways; you may need to trial a couple of recording methods in order to assess what is best for you.

- *Note taking* - this can range from highly structured to open and interpretive.
- *Audio recording* - audio recording allows you to preserve raw data for review at a later date.
- *Video taping* - video taping offers the added bonus of being able to record visual cues, but is more intrusive; is prone to more technical difficulties; and can generate data that is hard to analyse.

# Observation

❖ ***Observation:*** A systematic method of data collection that relies on a researcher's ability to gather data through their senses

❖ ***Observe:*** To notice using a full range of appropriate senses. To see, hear, feel, taste, and smell

# What You See
# Isn't Always What You Get

❖Observation provides the opportunity for researchers to document actual behaviour rather than responses related to behaviour

❖However, the observed can act differently when surveilled; and researchers' observations are likely to be biased by their own worldviews

# Types of Observation

❖In conducting observations, researchers can be anything from removed to immersed:

➢ ***Non-participant:*** In this role, the researcher does not become, nor aims to become an integral part of the system or community they are observing

➢ ***Participant:*** In this role, the researcher is, or becomes, a part of the team, community, or cultural group they are observing

# Types of Observation

❖ Researchers need to carefully consider the advantages and disadvantages of full disclosure:

➢ ***Candid/Overt:*** The researcher offers full disclosure of the nature of their study and the role the observations will play in their research

➢ ***Covert:*** Can be non-participant, i.e.) watching pedestrian behaviour at an intersection, or watching interactions at a school playground. But they can also be participatory. This involves researchers going 'undercover' in an attempt to get a real sense of a situation, context, or phenomenon

# Types of Observation

❖ Observational techniques can range from highly structured to unstructured.

> ➤ *Structured:* Highly systematic and often rely on predetermined criteria related to the people, events, practices, issues, behaviours, actions, situations, and phenomena being observed.

> ➤ *Semi-structured:* Observers generally use some manner of observation schedule or checklist to organize observations, but also attempt to observe and record the unplanned and/or the unexpected.

> ➤ *Unstructured:* Observers attempt to observe and record data without predetermined criteria.

# The Observation Process

❖The collection of credible data through observation requires:

  ➢thorough planning

  ➢careful observation

  ➢thoughtful recording

  ➢reflexive review

  ➢considered refinements

  ➢appropriate analysis

# Observation

❖ We do not all take in or perceive the world in the same way. In fact, two observers in the same situation will take in quite different sensory inputs

❖ It is exceptionally difficult for researchers to be objective. Our worldviews are a part of how we understand and make sense of the world, and how we might go about observing it

❖ Another difficulty in unbiased reflecting is your own expectations. Put simply, you are more likely to see the things you expect to see, and hear the things you want to hear

# Document Analysis

- ***Document analysis:*** Collection, review, interrogation, and analysis of various forms of text as a primary source of research data

# From Documents to Data

❖Two techniques for drawing data from a document are:

➢*The Interview* :  treating each document as a 'respondent' . Basically you 'ask' your document a question, and highlight the passages in the document that provides the answer.

➢*Noting occurrences* - quantifying the use of particular words, phrases, and concepts within a given document.

# Things to Consider

❖ All data collection methods are capable of gathering quantitative and qualitative data, although some may be better suited towards one task or the other

❖ There is no single data collection method that can guarantee credible data

❖ All data collection methods can be consciously manipulated

❖ All data collection methods can be 'contaminated' by unrecognized bias

❖ All data collection methods require conscious deliberation on the part of the researcher to ensure credibility

# References

➢ Uwyo.edu, (2015). [online] Available at: http://www.uwyo.edu/rgodby/econ%204360/data%20colle ction%20methods.ppt [Accessed 2 Feb. 2015].

➢ Wweb.uta.edu, (2015). [online] Available at: http://wweb.uta.edu/insyopma/baker/STATISTICS/Keller7/ Keller%20PP%20slides-7/Chapter05.ppt [Accessed 2 Feb. 2015].