Contact

Tang Weijun
Shanghai Jiao Tong University

E-mail: wj.tang@yahoo.com.cn

URL: http://www.drtang.org/
Learning points

- To identify the full variety of secondary data that are available.
- To appreciate ways in which secondary data can be utilized to help to answer research question and to meet objectives.
- To understand the advantages and disadvantages of using secondary data in research projects.
- To use a range of techniques, including published guides and the Internet, to locate secondary data.
- To evaluate the suitability of secondary data for answering research question and meeting objectives in terms of coverage, validity, reliability and measurement bias.
- To apply the knowledge, skills and understanding gained to your own research project.
7.1 Introduction

- Secondary data include both raw data and published summaries.
- Most organizations collect and store a variety of data to support their operations.
- Some of these data are available only from the organization that produce them, and so access will need to be negotiated.
- Others are widely available in published from as well as on CD in university libraries and increasingly via the Internet.
7.2 Types of secondary data and use in research

- Secondary data include both quantitative and qualitative data, and they can be used in both descriptive and explanatory research.
- Types of secondary data
  - See figure 7.1 in page 190
7.2 Types of secondary data and use in research

- Documentary data
  - Documentary secondary data are often used in research projects that also use primary data collection methods.
  - Written documents
    - Notices, correspondence, minutes of meetings, reports to shareholders, diaries, transcripts of speeches and administrative and public records.
  - Non-written documents
    - Tape and video recordings, pictures, drawings, films and television programmes, DVD/CD.
7.2 Types of secondary data and use in research

- Survey-based secondary data
  - Survey-based secondary data refers usually to data collected by questionnaires that have already been analyzed for their original purpose.
  - It will have been collected through three distinct types of survey:
    - Censuses
    - Continuous/regular surveys
    - Ad hoc surveys
7.2 Types of secondary data and use in research

- Survey-based secondary data
  - Censuses are usually carried out by governments and are unique because, unlike surveys, participation is obligatory. Consequently, they provide very good coverage of the population surveyed.
  - Continuous and regular surveys are those surveys, excluding censuses, that are repeated over time. They include surveys where data are collected throughout the year, and those repeated at regular intervals.
7.2 Types of secondary data and use in research

- **Survey-based secondary data**
  - Census and continuous and regular survey data provide a useful resource with which to compare or set in context your own research findings.
  - Survey secondary data may be available in sufficient detail to provide the main data set from which to answer your research question and to meet your objectives. Alternatively, they may be the only way in which you can obtain the required data.
7.2 Types of secondary data and use in research

- Survey-based secondary data
  - Ad hoc survey are usually one-off surveys and are far more specific in their subject matter. They include data from questionnaires that have been undertaken by independent researcher as well as surveys undertaken by organizations and governments. You will probably find it more difficult to discover relevant surveys.
7.2 Types of secondary data and use in research

- Multiple-source secondary data
  - Multiple-source secondary data can be based entirely on documentary or on survey data, or can be an amalgam or the two.
  - The key factor is that different data sets have been combined to form another data set prior to your accessing the data.
7.2 Types of secondary data and use in research

- **Multiple-source secondary data**
  - The way which a multiple-source data set has been compiled will dictate the sorts of research question or objectives with which you can use it.
    - Extract and combine selected comparable variables from a number of surveys or from the same survey that has been repeated a number of times to provide a time-series of data.
    - Compile for the same population over time using a series of “snap-shots” to form cohort studies.
    - For the data have the same geographical basis, to form area-based data sets.
7.3 Locating secondary data

- Finding relevant secondary data requires detective work:
  - Establishing that the sorts of data you require are likely to be available as secondary data.
  - Locating the precise data you require.
7.3 Locating secondary data

- The availability of secondary data
  - Books and journal articles on your chosen topic
  - Reference for unpublished and documentary
  - Tertiary literature such as indexes and catalogues
  - On-line indexes and catalogues
  - Informal discussions
7.3 Locating secondary data

- Finding secondary data
  - Precise references are often given in published guides, and a full reference should exist.
  - Data that are held by organizations are more difficult to locate.
  - Data on the Internet can be located using site guides.
  - Once you have located a possible secondary data set, you need to be certain that it will meet your needs.
7.4 Advantages and disadvantages of secondary data

- **Advantages:**
  - May have fewer resource requirements
  - Unobtrusive
  - Longitudinal studies may be feasible
  - Can provide comparative and contextual data
  - Can result in unforeseen discoveries
  - Permanence of data
7.4 Advantages and disadvantages of secondary data

- Disadvantages:
  - May be collected for a purpose that does not match your need
  - Access may be difficult or costly
  - Aggregations and definitions may be unsuitable
  - No real control over data quality
  - Initial purpose may affect how data are presented
7.5 Evaluating secondary data sources

- Secondary data must be viewed with the same caution as any primary data that you collect. You need to be sure that:
  - They will enable you to answer your research question and to meet your objectives.
  - The benefits associated with their use will be greater than the costs.
  - You will be allowed access to the data.
7.5 Evaluating secondary data sources

- If you are using secondary data, you are at an advantage compared with researchers using primary data. Because the data already exist you can evaluate them prior to use.
  - *See figure 7.2 in page 205*
- Alongside this process you also need to consider the accessibility of the secondary data.
7.5 Evaluating secondary data sources

- Overall suitability
  - Measurement validity
    - Secondary data that fail to provide you with the information that you need to answer your research question or meet your objectives will result in invalid answers.
    - Often when you are using secondary survey data you will find that the measures used do not quite match those that you need.
    - There are no clear solutions to problems of measurement invalidity.
7.5 Evaluating secondary data sources

- Overall suitability
  - Coverage and unmeasured variables
    - You need to be sure that the secondary data cover the population about which you need data, for the time period you need, and contain data variables that will enable you to answer your research question and to meet your objectives.
    - Ensuring that unwanted data are or can be excluded.
    - Ensuring that sufficient data remain for analyses to be undertaken once unwanted data have been excluded.
7.5 Evaluating secondary data sources

- Precise suitability
  - Reliability and validity
    - You will probably find the validity of documentary data such as organizations’ records more difficult to assess.
    - For all secondary data a detailed assessment of the validity and reliability will involve you in an assessment of the method or methods used to collect the data.
    - For some documentary sources, it is unlikely that there will be a formal methodology describing how the data were collected.
7.5 Evaluating secondary data sources

- Precise suitability
  - Reliability and validity
    - The validity and reliability of collection methods for survey data will be easier to assess where you have a clear explanation of the methodology used to collect the data.
    - Where data have been complied, as in a report, you need to pay careful attention to how these data were analyzed and how the results are reported.
7.5 Evaluating secondary data sources

- Precise suitability
  - Measurement bias
    - Measurement bias can occur for two reasons:
      - Deliberate or intentional distortion of data.
      - Changes in the way data are collected.
7.5 Evaluating secondary data sources

- Precise suitability
  - Measurement bias
    - Deliberate distortion occurs when data are recorded inaccurately on purpose, and is most common for secondary data sources such as organizational records.
    - Other distortion may be deliberate but not intended for any advantage.
    - Unfortunately, measurement bias resulting from deliberate distortion is difficult to detect.
    - Changes in the way in which data were collected can also introduce changes in measurement bias.
7.5 Evaluating secondary data sources

- Precise suitability
  - Cost band benefits
    - The final criterion for assessing secondary data is a comparison of the costs of acquiring them with the benefits they will bring.
    - Cost include both time and financial resources that you will need to devote to obtaining the data.
    - Benefits from data can be assessed in terms of the extent to which they will enable you to answer your research question and meet your objectives.
7.6 Summary

- Data that have already been collected for some other purpose, perhaps processed and subsequently stored, are termed secondary data. There are three main types of secondary data: documentary, survey, and those from multiple sources.
- Most research projects require some combination of secondary and primary data to answer your research question and to meet your objectives. You can use secondary data in a variety of ways.
7.6 Summary

- Any secondary data you use will have been collected for a specific purpose. This purpose may not match that of your research. In addition, the secondary data are likely to be less current than any data you collect yourself.
- Finding the secondary data you require is a matter of detective work.
- Once located you must assess secondary data sources to ensure their overall suitability for your research question and objective.
7.6 Summary

- You must also evaluate the precise suitability of the secondary data. Your evaluation should include both reliability and any likely measurement bias. You can then make a judgement on the basis of the costs and benefits of using the data in comparison with alternative sources.

- When assessing costs and benefits you need to be mindful that secondary data that are not completely reliable and contain some bias are better than no data at all if they enable you partially to answer your research question and to meet your objectives.
The End

😊

THANK YOU!

© 2008 TANG Weijun
Shanghai Jiao Tong University