

Data Analysis and Dissemination

Module 10C

Seong-Jin Yeon, Ph.D.

Director of International Strategic Research Center

Korean Institute of Criminology

Key Messages

- Data analysts must understand survey goals, the data, the audience and the issues being addressed.
- Analysis by statistical agencies should be objective and not conducted to promote specific policies.
- An overriding principle in analysis and dissemination is protecting respondent confidentiality.
- Varied dissemination formats aimed at different audiences can broaden the utility and use of the survey data.
- Presentation of survey results should be relevant, objective, accurate and transparent.

Data analysis

The process in which data are transformed into information to answer the central questions that the survey was designed to address.

Initial analysis

The first examination of data, usually by a research institute or the NSO, Provides basic information from the survey, addressing the goals for which the survey was designed to achieve.

Secondary Analysis

Conducted at a later stage either by the NSO or outside researchers. Can examine broad ranges of issues related to victimization and the operation of the criminal justice system.

Fundamental principles of analysis

Objectivity

All findings must be supported by the data. Findings should not be directed towards or against supporting any specific policy or political agenda.

Understand the survey Goals

Analyses should be directed at answering the specific questions addressed in the survey's goals.

Understand the data

The analyst must know the survey's methodology and concepts used to collect the data, the structure of the questionnaire and the individual questions, and the strengths and limitations of the data.

Understand the audience and the issues

This is necessary so that the information can be presented in the most useful format and domain.

Questions to consider before beginning data analysis:

What issues am I addressing?

For example:

- Prevalence of crime
- Characteristics of victims
- Dark Figure
- Perceptions

What are the limitations of my data?

For example:

- Is subnational data possible?
- Is the sample size big enough to measure rare events such as some crime types?

What analytical methods are appropriate?

For example:

- Descriptive
- Inferential

Descriptive statistics
Descriptive statistics are used to describe the basic features of a collection of data. The most common measures of descriptive statistics are the mean, median, mode, range, and standard deviation. These measures provide a summary of the data and allow for comparison with other data sets.

Inferential statistics
Inferential statistics are used to make inferences about a population based on a sample of data. The most common measures of inferential statistics are the confidence interval and the p-value. These measures allow for the estimation of the true value of a population parameter and the testing of hypotheses.

What results do I want to highlight?

For example:

- Regional patterns
- Demographic differences
- Impact of crime on victims
- Trends over time (short or long term)

What are my key findings?

- Which findings are related to the purposes for which the survey was conducted?
- Which findings may have the greatest impact on policy? (Note: this should be determined while maintaining objectivity.)
- Which findings will invoke the greatest public interest?



What issues am I addressing?

For example:

- Prevalence of crime
- Characteristics of victims
- Dark Figure
- Perceptions

What are the limitations of my data?

For example:

- Is subnational data possible?
- Is the sample size is big enough to measure rare events such as some crime types?

What analytical methods are appropriate?

For example:

- Descriptive
- Inferential

Descriptive statistics

Descriptive statistics are results relating to the population at the time the data were collected, and are generally presented as counts, percents, and rates.

Examples of descriptive statistics:

- About 50% of all violent victimizations and nearly 40% of property crimes [in the United States] were reported to police in 2010 (Source: U.S. Criminal Victimization 2010)
- En 2008 cerca del 92% de las víctimas [in Mexico] sufrieron el delito en la misma entidad donde residen, 5.5% en un estado diferente a donde viven y 2.5% sufrió delitos tanto en su estado como en algún. (Source: Victimization incidencia y cifra negra en Mexico. Analisis de la ENSI-6 p. 24)
- It is estimated that in the 12 months prior to interview in 2010-11, of the 17.7 million people aged 15 years and over in Australia 486,500 (2.7%) were victims of at least one physical assault. (Source: <http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4530.0Main+Features32010-11>)

Inferential statistics

Inferential statistics examine the relationship between variables, often using regression coefficients to describe how degrees of change in one variable impact changes in other variables.

Examples of studies that use inferential statistics to explore issues related to victimization include:

- How neighborhood characteristics such as levels of poverty impact residents becoming victims of crime.
- The relationship between victim resistance and injury during robberies.
- How various variables (demographic, lifestyle, geographic, etc.) impact risks of becoming a victim of specific types of crime.

Descriptive statistics

Descriptive statistics are results relating to the population at the time the data were collected, and are generally presented as counts, percents, and rates.

Examples of descriptive statistics:

- About 50% of all violent victimizations and nearly 40% of property crimes [in the United States] were reported to police in 2010 (Source: U.S. Criminal Victimization 2010)
- It is estimated that in the 12 months prior to interview in 2010-11, of the 17.7 million people aged 15 years and over in Australia 486,500 (2.7%) were victims of at least one physical assault.

(Source: [http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4530.0
Main+Features32_010-11](http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4530.0>Main+Features32_010-11))

Inferential statistics

Inferential statistics examine the relationship between variables, often using regression coefficients to describe how degrees of change in one variable impact changes in other variables.

Examples of studies that use inferential statistics to explore issues related to victimization include:

- How neighborhood characteristics such as levels of poverty impact residents becoming victims of crime.
- The relationship between victim resistance and injury during robberies.
- How various variables (demographic, lifestyle, geographic, etc,) impact risks of becoming a victim of specific types of crime.

What results do I want to highlight?

For example:

- Regional patterns
- Demographic differences
- Impact of crime on victims
- Trends over time (short or long term)

What are my key findings?

For example:

- Which findings are related to the purposes for which the survey was conducted?
- Which findings may have the greatest impact on policy?
(Note: this should be determined while maintaining objectivity.)
- Which findings will invoke the greatest public interest?

Disseminating survey results

Once data have been analyzed, the results must be interpreted and presented to users.

Presentation of survey results should be relevant, objective, accurate and transparent.

Dissemination formats can take many forms:

- Analytical and statistical reports
- Fact sheets
- Press releases and press conferences
- Journal articles
- Websites
- Online data repositories and analysis tools
- Public use data files

The form of presentation should fit the audience and the release's purpose

Victimization survey dissemination products

The format, substance and publication medium chosen for release should be selected to meet the needs of the data users.

For the general public, shorter reports containing easy to read tables and graphics will ensure that the information will be understood.

Publications intended for use by government officials and policy makers should present the data in a direct way that highlights key findings. Such releases should include the strengths and weaknesses of the data so that they will be cited correctly if used to support policies.

Comparisons of victimization survey estimates and administrative data

Reports containing victimization survey estimates often present data from administrative records for comparison. In some jurisdictions the same agency is responsible for both data series.

Reports containing both survey and police data should stress the complementarity of the two series, and also discuss their differences in crime coverage and definition.

In focusing on the utility of survey victimization estimates it is useful to stress survey data relating to crimes not reported to police, and the detailed information about victims and offenses not available from police based estimates.



Conducting data analysis is like drinking fine wine. It is important to swirl and sniff the wine, to unpack the complex bouquet and to appreciate the experience. Gulping the wine doesn't work.

Daniel B. Wright

THANK YOU

