Survey Design 101

Part II: Analyzing and Interpreting Survey Findings

Okan Bulut, Ph.D.
Centre for Research in Applied Measurement and Evaluation
University of Alberta
Hello!

I am Okan Bulut

Associate Professor in Measurement, Evaluation, and Data Science (MEDS) & Program Coordinator

You can find me (virtually) at: bulut@ualberta.ca

@drokanbulut, www.okanbulut.com
Agenda

1. Analyzing survey data
2. Interpreting survey findings
Analyzing survey data

Data analysis options based on the type of survey data
How to analyze survey data

- What type of variables (quantitative or qualitative) do you have in your survey data?

- What are the possible options for analyzing survey data?
  - Statistical and qualitative data analysis
  - Software programs (open-access vs. commercial)

- Who is the target audience for the findings of your survey?

- How do you plan to present your findings?
  - Thesis/dissertation
  - Publications, technical reports
  - Posters and presentations
Types of variables

Surveys typically yield two types of data:

• **Quantitative data** (*less common*)
  - Age, income, height, weight
  - Time-related variables (e.g., years of experience)
  - Counts of things (e.g., children, books, credits earned)

• **Qualitative data** (*more common*)
  - Nominal categories (gender, race, province of residence)
  - Ordered categorical (level of satisfaction or agreement, frequency of a behaviour)
  - Descriptive comments (e.g., open-ended questions)
Descriptive statistics

Some descriptive statistics that we can report for survey data include:

1. **For quantitative variables:**
   - Mean, median, standard deviation, minimum, and maximum

2. **For qualitative variables (either ordinal or nominal):**
   - Nominal variables: Frequencies and proportions (i.e., percentages)
   - Ordinal variables: Frequencies, proportions, median, and mode
Avoid: Averaging ordinal variables

Ordinal variables such as Likert-scale questions should not be averaged or summed.
Item and scale analysis

1. Check response frequencies
   - Response categories (enough utilization of each response option)
   - Skewness in the responses (no extreme skewness is desired)
   - Levels of missing data (should be less than 10%)

Q1: I enjoy learning mathematics.

Response options are:
- 0 indicating "I strongly disagree"
- 1 indicating "I disagree"
- 2 indicating "I agree"
- 3 indicating "I strongly agree"

Frequency (%)

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>(Missing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>15%</td>
<td>17%</td>
<td>35%</td>
<td>23%</td>
<td>10%</td>
</tr>
</tbody>
</table>
2. Check the alignment among the questions

Questions focusing on “enjoying math” on the survey

Q1: I enjoy learning mathematics.
Q2: I wish have not to study Math. (reverse-coded)
Q3: Mathematics is boring. (reverse-coded)
Q4: I learn interesting things in mathematics class.
Q5: I like mathematics.
Q6: I think it’s important to do well in mathematics.
3. Check scale reliability (for a group of questions)

Questions focusing on “enjoying math” on the survey

Q1: I enjoy learning mathematics.
Q2: I wish have not to study Math. (reverse-coded)
Q3: Mathematics is boring. (reverse-coded)
Q4: I learn interesting things in mathematics class.
Q5: I like mathematics.
Q6: I think it’s important to do well in mathematics.
Analyzing qualitative responses

◉ Thematic analysis of open-ended responses
  ○ Generating initial codes or labels
  ○ Searching for themes or common patterns
  ○ Defining and naming each theme

◉ Word clouds using open-ended responses

◉ Text mining analysis
  ○ Sentiment analysis to extract emotions from the text
  ○ Topic modeling to create themes automatically
Statistical software programs

- Open-access (free) programs
  - jamovi - https://www.jamovi.org/
  - JASP - https://jasp-stats.org/
  - PSPP - https://www.gnu.org/software/pspp/
  - R (no GUI) - https://cran.r-project.org/

- Commercial programs
  - Microsoft Excel
  - IBM SPSS Statistics
  - Minitab
Important: Qualitative software programs facilitate data storage, coding, retrieval, comparing, and linking; but they do not analyze data (Patton, 2015, p. 529)

- Open-access programs
  - Compendium – for mapping and management of ideas and arguments
  - KH Coder – for content analysis and text mining
  - R and Python (no GUI)

- Commercial programs
  - MAXQDA, ATLAS.ti, and NVIVO
    - Check out UAlberta On the Hub for discounted prices
  - discovertext for analyzing textual data
Interpreting survey findings

How to interpret and report survey findings
Presenting survey findings

Tables

Figures
Tables vs. Figures

- **Prefer tables if:**
  - The table will be used to look up *individual* values
  - Precise information is required
  - Both *summary* and *original* values are needed

- **Prefer figures if:**
  - The goal is to reveal relationships among whole sets of values
  - The relationships between two or more variables are the main priority
Table 1. **Median** scores for items on the **Social Engagement Scale** by students’ pass/fail status in mathematics.

<table>
<thead>
<tr>
<th>As a 9th grade student, I …</th>
<th>Fail</th>
<th>Pass</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>care if I finish high school.</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>know my way around the school.</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>have many friends at school.</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>know what my teachers think about me.</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>ignore what other students think about me.*</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>come to school only because my friends are here</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

**Note**: Students failing mathematics $n = 31$; students passing mathematics $n = 147$. Response scale for the questions: 1 = Strongly disagree; 2 = Disagree; 3 = Agree; and 4 = Strongly Agree.

* The question uses **negative wording**, and a low median indicates positive student behaviour.
Avoid: Tables with a single row

<table>
<thead>
<tr>
<th></th>
<th>Telephone Survey</th>
<th>City Insight Panel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Satisfied (4,5 ratings)</td>
<td></td>
</tr>
<tr>
<td>2014 n=400</td>
<td>77%</td>
<td>75%</td>
</tr>
<tr>
<td>2015 n=400</td>
<td>80%</td>
<td></td>
</tr>
<tr>
<td>2014 n=1,126</td>
<td></td>
<td>74%</td>
</tr>
<tr>
<td>2015 n=1,630</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall satisfaction with living in the City of Edmonton

Source: 2015 Edmonton Citizen Perception Survey
Creating effective figures

- Figures can be very effective for summarizing survey findings.

- Consider the following elements
  - Colours
  - Shapes
  - Size
  - Font type, size, and colour

- Check out my workshop notes on visualizing survey items
Avoid: Complex figures
Less is more...

Remove to improve (the data-ink ratio)

Source: Darkhorse Analytics – Data looks better naked
Bar charts

### PISA 2009 Results (Source: OECD)

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I find it hard to finish books.</td>
<td>9%</td>
<td>17%</td>
<td>34%</td>
<td>40%</td>
</tr>
<tr>
<td>I like talking about books with other people.</td>
<td>11%</td>
<td>28%</td>
<td>28%</td>
<td>33%</td>
</tr>
<tr>
<td>Reading is one of my favorite hobbies.</td>
<td>13%</td>
<td>24%</td>
<td>29%</td>
<td>34%</td>
</tr>
<tr>
<td>I read only if I have to.</td>
<td>14%</td>
<td>24%</td>
<td>26%</td>
<td>36%</td>
</tr>
</tbody>
</table>

**Note:** The data come from a random sample of Canadian students who participated in PISA 2009.
Bar charts

PISA 2009 Results (Source: OECD)

I find it hard to finish books.
- Strongly disagree: 34%
- Disagree: 40%
- Agree: 17%
- Strongly agree: 9%

I like talking about books with other people.
- Strongly disagree: 28%
- Disagree: 28%
- Agree: 33%
- Strongly agree: 11%

Reading is one of my favorite hobbies.
- Strongly disagree: 29%
- Disagree: 34%
- Agree: 24%
- Strongly agree: 13%

I read only if I have to.
- Strongly disagree: 26%
- Disagree: 36%
- Agree: 24%
- Strongly agree: 14%

Note: The data come from a random sample of Canadian students who participated in PISA 2009.
“Some” tools to make great figures

◉ Microsoft Excel:
  ○ Bar charts, pie charts, scatter plots, line charts, etc.
  ○ Excel gives publication-quality graphs (a bit tedious...)
  ○ [https://www.youtube.com/watch?v=d7JgmQCLgNg](https://www.youtube.com/watch?v=d7JgmQCLgNg)

◉ jamovi
  ○ Bar charts, scatterplots, boxplots, correlation matrix plots, etc.
  ○ Not very customizable but high-quality graphs
  ○ [https://www.youtube.com/watch?v=NoG-YJcTtx8](https://www.youtube.com/watch?v=NoG-YJcTtx8)

◉ Websites to create graphs
  ○ RAWGraphs ([https://rawgraphs.io/](https://rawgraphs.io/)): Free and easy to use
  ○ Datawrapper ([https://www.datawrapper.de/](https://www.datawrapper.de/)): Free plan; charts, tables, and maps
  ○ Canva Graph Maker ([https://www.canva.com/graphs/](https://www.canva.com/graphs/)): Beautiful infographics; values must be entered manually.
Thanks!

Any questions?