

# Outcomes for Success!

**Session #3: Data Management, Analysis & Reporting**

**King County Library System**  
**October 5, 2010**

Presented by  
**Organizational Research Services**  
Seattle, WA  
206.728.0474  
[www.organizationalresearch.com](http://www.organizationalresearch.com)

- Improved ability to manage your data
- Increased knowledge of techniques for data analysis and basic statistics
- Increased ability to interpret evaluation data
- Increased ability to communicate findings with stakeholders

***How do we collect, store and use data  
about a program?***

## ***Allow your organization to:***

- Easily use the information collected on program participants
- Add information to and extract it from a database
- Minimize the time and expense associated with the process of collecting and analyzing outcome data

- Software and Data Management Systems
- Client ID Considerations
- Coding Survey Questions
- Codebooks
- Quality Checks

- **Microsoft Excel**
- Microsoft Access
- Survey Monkey
- SPSS
- Others

## ***Consider the following:***

- Learning Curve
- System issues: What size data collection and storage efforts can our system handle?
- Expertise: What is the level of expertise in my organization?
- Staff Turnover
- What systems already exist?

## **Exercise**

# Setting Up a Data Management System

- ROWS represent the cases or participants
- COLUMNS represent the “variables” collected

| ID | Variable 1 | Variable 2 | Variable 3 | Variable 4 | Variable 5 |
|----|------------|------------|------------|------------|------------|
| 1  | Data       | Data       | Data       | Data       | Data       |
| 2  | .          | .          | .          | .          | .          |
| 3  | .          | .          | .          | .          | .          |

# Advantages of Creating One Combined Data File

- Eliminates the need to match data sets and merge files
- Provides all data on any participant for a particular set of outcomes in one data file
- Reduces the number of data sets and spreadsheets

**EXAMPLE:** create one data set where baseline and follow-up data are linked by a

Participant's ID Number

| ID | Baseline   |            |            | Follow-Up  |            |            |
|----|------------|------------|------------|------------|------------|------------|
|    | Question 1 | Question 2 | Question 3 | Question 1 | Question 2 | Question 3 |
| 1  | 3          | 2          | 2          | 4          | 2          | 4          |
| 2  | 2          | 3          | 3          | 3          | 3          | 3          |
| 3  | 4          | 1          | 4          | 5          | 2          | 5          |

- Track participants over time
- Administer a certain number of tools
- Confidentiality and anonymity
- Avoid data matching errors associated with names

- Number surveys
- Personal identifiers
- Complex ID numbers:
  - Useful for anonymity
  - Use information provided by participant

- Codebooks represent a **formal set of instructions** to guide the data entry and coding process
- Codebooks help **multiple staff members understand** the data entry and analysis process

| Questions  | Variable Names                 | Possible Answers or Response Codes   |
|--|--------------------------------|--|
| ID   | ID                             | Enter number assigned to each respondent   |
| Survey date  | DATE                           | Enter date mm/dd/year  |
| Gender   | GENDER                         | 1 = Male<br>2 = Female   |
| Confidence about ability to study each night   | CONF                           | 1 = strongly disagree<br>2 = disagree<br>3 = neither agree nor disagree<br>4 = agree<br>5 = strongly agree |
| Materials that have helped study skills over past six months: workbook, CD-ROM, Tutoring lessons | WORKBOOK<br>CD-ROM<br>TUTORING | For each option enter:<br>0 = if box not checked<br>1 = if box checked                                     |
| Rate of classroom participation  | PARTICIPATES                   | 1 = never<br>2 = not very often<br>3 = somewhat often<br>4 = very often                                    |

## ***Two common types of items used in evaluation tools:***

- **Open-ended** questions. Responses are provided by respondents.
- **Closed-ended** questions. Pre-determined answers from which participants choose (*e.g., scales, check-off boxes*).

## ***Close-ended Questions—Scales:***

- Responses are pre-determined and rank ordered along continuum
- Respondents only choose one category
- Numeric scale is used to represent order of responses for analysis
- Often a 'middle' category like 'neither agree nor disagree' or 'unsure'

## *Example:*

### *Scale with coding in parentheses*

- Very dissatisfied (1)
- Somewhat dissatisfied (2)
- Neither satisfied nor dissatisfied (3)
- Somewhat satisfied (4)
- Very satisfied (5)

## ***Close-ended Question—Check-Off Boxes:***

- Only one box can be checked
- More than one box can be checked

Response 1

**Response 2**

Response 3

Response 4

**OR**

Response 1

**Response 2**

**Response 3**

**Response 4**

## ***Check-off Boxes:***

### ***Criteria when ONLY one box can be checked***

- Numeric value is assigned to each of the possible responses
- Options are mutually exclusive (*i.e., participant's response can not fall into two or more categories*)
- Options are exhaustive (*e.g., cover the full range of possible responses*)

## *Example:* Coding in parentheses

### What is your age?

- 16-18 years old (1)
- 19-20 years old (2)
- 21-24 years old (3)
- 25-29 years old (4)
- 30-32 years old (5)

## ***Check-Off Boxes:***

### ***Criteria when MULTIPLE BOXES can be checked***

- **Any number of responses** may be checked
- Each response category is a **separate question**
- **Different variables for different response categories**

## *Example: Coding in parentheses\**

### How do you communicate with your child(ren)?

*(Please check all of the responses that apply to you and your family.)*

- Sign language **(0 not checked, 1 checked)**
- Home signs **(0 not checked, 1 checked)**
- Speech **(0 not checked, 1 checked)**
- Paper and pencil **(0 not checked, 1 checked)**
- Gestures **(0 not checked, 1 checked)**
- Other **(0 not checked, 1 checked)**

| Communicate with Children |               |            |        |
|---------------------------|---------------|------------|--------|
| ID                        | Sign Language | Home signs | Speech |
| 1                         | 1             | 0          | 1      |
| 2                         | 0             | 1          | 1      |
| 3                         | 0             | 1          | 0      |
| 4                         | 1             | 1          | 1      |

**\***

Each response is considered a unique variable.

***Use the following values when data are either missing or skipped:***

- 9 = Missing data
- 8 = Not applicable

\* Please note that one SHOULD NOT use zero (0) as a code for missing values because it is a valid response and thus may lead to inaccurate calculations.

## ***Once your data are entered:***

- Conduct a scan for obvious errors
- Have another person check 10% of the surveys to ensure accuracy of coding and data entry
- Discuss and resolve any errors that are found

***How do we analyze data once they have been coded and entered?***

- ***Frequencies:*** summaries of the number or percent of observations in each response category
- ***Averages:*** mean of responses
- ***Median:*** midpoint of all cases
- ***Mode:*** most frequently occurring response
- ***Standard Deviation:*** variation of distribution
- ***Cross-tabulations:*** summaries of frequency distributions across different subgroups or levels of a second variable

- **Frequency** (count) – number of individuals in a response category
- **Percentage** – percent of total individuals in a response category

# Frequency Distribution For Satisfaction Item

| SATISFACTION                       | Frequency | Percentage (%) |
|------------------------------------|-----------|----------------|
| Very Satisfied                     | 6         | 30%            |
| Somewhat Satisfied                 | 4         | 20%            |
| Neither Satisfied nor Dissatisfied | 5         | 25%            |
| Somewhat Dissatisfied              | 3         | 15%            |
| Very Dissatisfied                  | 2         | 10%            |
| <b>OVERALL TOTAL</b>               | <b>20</b> | <b>100%</b>    |

## ***Frequencies:***

***‘Eyeball’ these data. What do they tell you about this program’s settings?***

| PROGRAM SETTING        | N          | %            |
|------------------------|------------|--------------|
| School                 | 144        | 61.5         |
| Community-Based Agency | 38         | 16.2         |
| Faith-Based Agency     | 35         | 15.0         |
| Other                  | 17         | 7.3          |
| <b>TOTAL</b>           | <b>234</b> | <b>100.0</b> |

**Mean**

**Median**

**Mode**

***Consider these data:***

***What do they tell you about the grades served?***

| Grade  | N    | %     |
|--------|------|-------|
| 1      | 29   | 12.8  |
| 2      | 48   | 21.2  |
| 3      | 47   | 20.8  |
| 4      | 40   | 17.7  |
| 5      | 28   | 12.4  |
| 6      | 10   | 4.4   |
| 7      | 4    | 1.8   |
| 8      | 12   | 5.3   |
| 9      | 8    | 3.5   |
| Total  | 226  | 100.0 |
| Mean   | 3.64 |       |
| Median | 3    |       |
| Mode   | 2    |       |

- Compare distributions across subgroups
- Two-way or Contingency tables
- Cells indicate number or percentage of respondents

| SATISFACTION                          | Males      | Females    | ALL        |
|---------------------------------------|------------|------------|------------|
| Very Satisfied                        | 1<br>(10%) | 5<br>(50%) | 6<br>(30%) |
| Somewhat Satisfied                    | 3<br>(30%) | 1<br>(10%) | 4<br>(20%) |
| Neither Satisfied nor<br>Dissatisfied | 4<br>(40%) | 1<br>(10%) | 5<br>(50%) |
| Somewhat Dissatisfied                 | 2<br>(20%) | 1<br>(10%) | 3<br>(30%) |
| Very Dissatisfied                     | 0<br>(0%)  | 2<br>(20%) | 2<br>(20%) |
| <b>OVERALL TOTAL</b>                  | <b>10</b>  | <b>10</b>  | <b>20</b>  |

## ***Cross-tabulations:***

*What do these data tell you about the location of each of the settings?*

| PROGRAM SETTING        | % E. Wash.<br>(N)  | % W. Wash.<br>(N) |
|------------------------|--------------------|-------------------|
| School                 | 63.9 (108)         | 55.4 (36)         |
| Community-Based Agency | 22.5 (38)          | 0.0 (0)           |
| Faith-Based Agency     | 13.6 (23)          | 18.5 (12)         |
| Other                  | 0.0 (0)            | 26.2 (17)         |
| <b>TOTAL</b>           | <b>100.0 (169)</b> | <b>100.0 (65)</b> |

## ***T-tests***

- Allow you to analyze differences between groups or over time. Test for statistically significant difference between mean values.
- Assesses probability of whether changes are due to chance (p value). Standard convention  $p < .05$  (probability that difference is due to chance is less than 5 percent).

- Independent Samples – comparison of mean values for one variable across subgroups  
*(e.g., Males vs. Females)*
- Paired Samples – comparison of mean values on one variable over time for the same participants  
*(e.g., Pre vs. Post)*

|   | Participants<br>(n=48) | Non-Participants<br>(n=37) | p-value |
|---|------------------------|----------------------------|---------|
| Average GPA                               | 2.45                   | 1.95                       | .012*   |
| Skills Ranking (1-10)                     | 7.43                   | 5.85                       | .003*   |
| Self-Esteem Score<br>(1-100 scale)        | 64.5                   | 58.9                       | .098    |
| Average Discipline<br>Referrals Per Month | 1.87                   | 3.28                       | .001*   |

## ***Look at your data:***

- What findings are most interesting?
- What patterns do you see?
- What client characteristics might explain these patterns?
- What program strategies might explain these patterns?

***Open-Ended: The respondent provides own responses to questions with no predetermined set of responses.***

***The process of coding:***

- Identify “naturally occurring” categories
- Reduce number of categories
- Create rules and procedures
- Test coding/inter-coder reliability

## ***Sample of five specific responses from an open-ended survey question***

- “Organization XX should expand their hours of service”
- “The organization staff should gain more experience”
- “The organization should train their staff better”
- “The organization should give out more financial aid”
- “Organization XX should stay open later”

## ***Completed coding schema***

- Time/hours issues:
  - “Organization XX should stay open later”
  - “Organization XX should expand their hours of service”
- Staffing issues:
  - “The organization staff should gain more experience”
  - “The organization should train their staff better”
- Financial issues:
  - “The organization should give out more financial aid”

***How do we share our findings with stakeholders?***

## ***Consider your audiences:***

- **Internal stakeholders:** staff, participants, board
- **External stakeholders:** funders, partners, media

## ***Purposes:***

- Quality improvement (*i.e., share program modifications*)
- Marketing
- Accountability
- Celebrate successes
- Tell your story

## ***Methods of presenting findings visually***

- Bar and grouped bar charts
- Pie charts
- Line graphs
- Scatterplots

- Report the frequency/percentage of responses in each category
  - If small N or non-representative sample, can share the range and prevalence of different themes (e.g., “most respondents said X”)
- Share illustrative quotes and examples for each category/theme

## ***Describe your program implementation.***

- To what extent did implementation go according to plan?
  - What resources did you have in place?
  - What activities did you do?
  - What were your outputs?

## ***Changes as a result of your program***

- What were your participant outcomes?
- Who changed?
- How do you know?

## ***Reflection:***

### ***Observations, Interpretations, Implications***

- What observations can you make about what was presented under previous questions?
- Do the data presented allow you to reach strong conclusions?
- What questions have been raised?
- What potential implications have been raised?

## ***Key 'To Do's':***

- Are outcomes and indicators clearly identified?
- Is the sample clearly identified?
- Are data sources and collection methods clearly described?
- Are data presented clearly and accurately?
- Are interpretations sensible, logical and clear?

- Describe findings accurately, honestly and simply
- If findings aren't optimal:
  - Focus on progress
  - Identify specific program improvement actions
  - Learn what's working in similar programs
  - Reassure audiences that you are aware of issues and are working to address them

## ***After the report, what next?***

- Consider external factors that may have impacted results
- Review program theory, tools, chosen outcomes, methodologies, participant and program characteristics
- Double check analysis, data entry, formulas, etc.
- Is further analysis needed?
- Are program modifications warranted?

## Organizational Research Services

[www.organizationalresearch.com](http://www.organizationalresearch.com)

206-728-0474

### Joelle Cook, MPA

[jcook@organizationalresearch.com](mailto:jcook@organizationalresearch.com), ext. 234

### Steve Mumford, MPA

[smumford@organizationalresearch.com](mailto:smumford@organizationalresearch.com), ext. 224

### Emilee Quinn, MPH

[equinn@organizationalresearch.com](mailto:equinn@organizationalresearch.com), ext. 227