

Basic Data Analysis

MER

Midwest
Evaluation and
Research, LLC

Objectives

- Understand the definition and purpose of data analysis
- Understand the core steps for data analysis (clean, organize, and analysis)
- Understand three basic methods of data analysis (frequencies, centers, and cross tabulation)
- Understand how to interpret and present data

Overview

- What is it?
- Data Analysis Plan
- Types of Data
- Quantitative Analysis
- Qualitative Analysis
- Interpretation
- Reflection
- Presentation/Visualization

What is it?

MER

Midwest
Evaluation and
Research, LLC

Data Analysis

- Turning raw data into useful information
- Provide answers to programmatic and/or research questions
- The greatest amount and best quality data mean nothing if not properly analyzed—or if not analyzed at all

Data Analysis Steps

- Start with a plan
- Collect data
- Analyze
- Interpret
- Reflect
- Presentation/Visualization

Data Analysis Plan

MER

Midwest
Evaluation and
Research, LLC

Data Analysis Plan

- Purpose of the Analysis
- Key Questions
- Data Sources and Collection
- Analysis Method(s)
- Report and Present Results

Types of Data

Two Kinds of Data

- Quantitative
- Qualitative

Quantitative Data

- Data that is numerical, counted, or compared on a scale
 - Demographic data
 - Answers to closed-ended survey items
 - Attendance data
 - Scores on standardized instruments

Quantitative Data Collection Methods

- Surveys / Questionnaires
- Interviews
- Observations

Qualitative Data

- Textual data
- Photographs
- Audio recordings
- Video recordings

Qualitative Data Collection Methods

- Focus groups
- Interviews
- Open-ended survey items
- Document review
- Notes from observations

Working Assumptions

MER

Midwest
Evaluation and
Research, LLC

Working Assumptions

- Data is clean
- Check errors in a systematic way
- Codebook
 - Defines and lays out the structure of each variable

Quantitative Analysis

MER

Midwest
Evaluation and
Research, LLC

Quantitative Analysis

- Data tabulation
- Descriptive data
- Data disaggregation

Data Tabulation

- This will help you determine:
 - If scores are entered correctly
 - If scores are high or low
 - How many are in each category
 - The spread of the scores

Data Tabulation

- Frequency distributions
 - Counts of the number of responses to a question or occurrences
- Percent distributions
 - Proportion of participants who are represented within each variable

Frequency Distribution

Trends

Variable	Frequency
Gender	Female – 70% Male – 30%
Race	African-Amer – 30% Caucasian – 54% Asian – 8% Native Amer – 8%

Variable	Frequency
Number of school events	< 5 = 39% 5 – 10 = 33% > 10 = 28%
Change in Knowledge	-2 – 2% -1 – 16% 0 – 16% 1 – 25% 2 – 18% 3 – 17% 4 – 6%

Frequency Distribution

Objectives

Variable	Frequency
Gender	Female – 70% Male – 30%
Race	African-Amer – 30% Caucasian – 54% Asian – 8% Native Amer – 8%

Variable	Frequency
Number of school events	< 5 = 39% 5 – 10 = 33% > 10 = 28%
Change in Knowledge	-2 – 2% -1 – 16% 0 – 16% 1 – 25% 2 – 18% 3 – 17% 4 – 6%

Frequency Distribution

Context

Variable	Frequency
Gender	Female – 70% Male – 30%
Race	African-Amer – 30% Caucasian – 54% Asian – 8% Native Amer – 8%

Variable	Frequency
Number of school events	< 5 = 39% 5 – 10 = 33% > 10 = 28%
Change in Knowledge	-2 – 2% -1 – 16% 0 – 16% 1 – 25% 2 – 18% 3 – 17% 4 – 6%

Descriptives

- Describes the program and its participants
- Does not define causality – tells you what, not why
- Common descriptives:
 - **Mode** - variable with highest frequency
 - **Median** - middle variable or score that splits the distribution of a variable into half
 - **Mean** - average score of a variable
 - **Minimum and maximum values** – the highest and lowest value for a particular variable

Frequency Distribution & Centers

Variable	Frequency	Center
Number of school events	< 5 = 39% 5 – 10 = 33% > 10 = 28%	Mode – 2 Median – 6 Mean – 6
Change in Knowledge	-2 – 2% -1 – 16% 0 – 16% 1 – 25% 2 – 18% 3 – 17% 4 – 6%	Mode – 1 Median – 1 Mean – 1

Disaggregating the Data

- Crosstabs allow you to disaggregate the data
 - across multiple categories
 - by subcategories within a variable

Cross Tabulation

		Gender		Total (n=100)
		Female (n=70)	Male (n=30)	
Race	African-America	12	18	30
	Caucasian	44	10	54
	Asian	7	1	8
	Native American	7	1	8
Number of School Events	Less than 5	26	13	39
	5 - 10	26	7	33
	Above 10	18	10	28

Qualitative Analysis

MER

Midwest
Evaluation and
Research, LLC

Qualitative Analysis

- Transcribe data (if audio taped)
- Read transcripts
- Highlight quotes and note why important
- Code quotes
- Sort quotes into coded groups (themes)
- Interpret patterns in quotes
- Describe these patterns

Data Analysis Tools

- Spreadsheet – simple to use, basic graphs
- Statistical packages, e.g. SPSS
- Qualitative data analysis tools, e.g. Nvivo

Interpretation

MER

Midwest
Evaluation and
Research, LLC

Interpreting Information



Interpreting Information



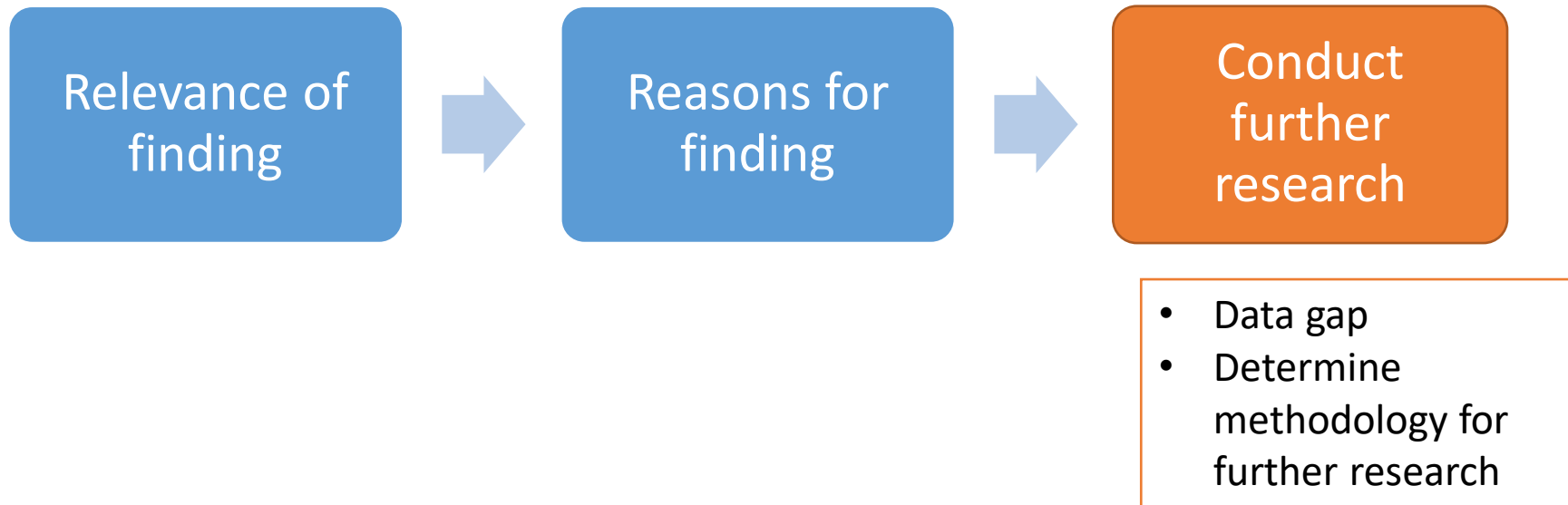
- Pair results to the key evaluation questions
- Revisit logic model and theory of change
- Does the results meet the target?
- How far from the target is it?
- How does it compare (to other time periods, other groups)?
- Are there any extreme highs and lows in the data?
- Did anything surprise you?
- Any interesting or confusing patterns and trends?

Interpreting Information



- Why? – Potential reasons for findings
- Program context – social, financial, political etc
- Other data sources
 - Situation analyses
 - Performance data

Interpreting Information




Reflection



MER

Midwest
Evaluation and
Research, LLC



Reflection

- What is important in the data?
- Why is it important?
- What can be learned from it?
- So what?
- What conclusions can we draw?
- What are our recommendations?
- What are the limitations of our analysis?

Presentation/Visualization

MER

Midwest
Evaluation and
Research, LLC

Summarizing Data

- Tables
 - Simplest way to summarize data
 - Data are presented as absolute numbers or percentages
- Charts and graphs
 - Visual representation of data
 - Data are presented as absolute numbers or percentages

Basic Guidance When Summarizing Data

- Ensure graphic has a title
- Label the components of your graphic
- Indicate source of data with date
- Provide number of observations ($n=xx$) as a reference point
- Add footnote if more information is needed

Charts & Graphs

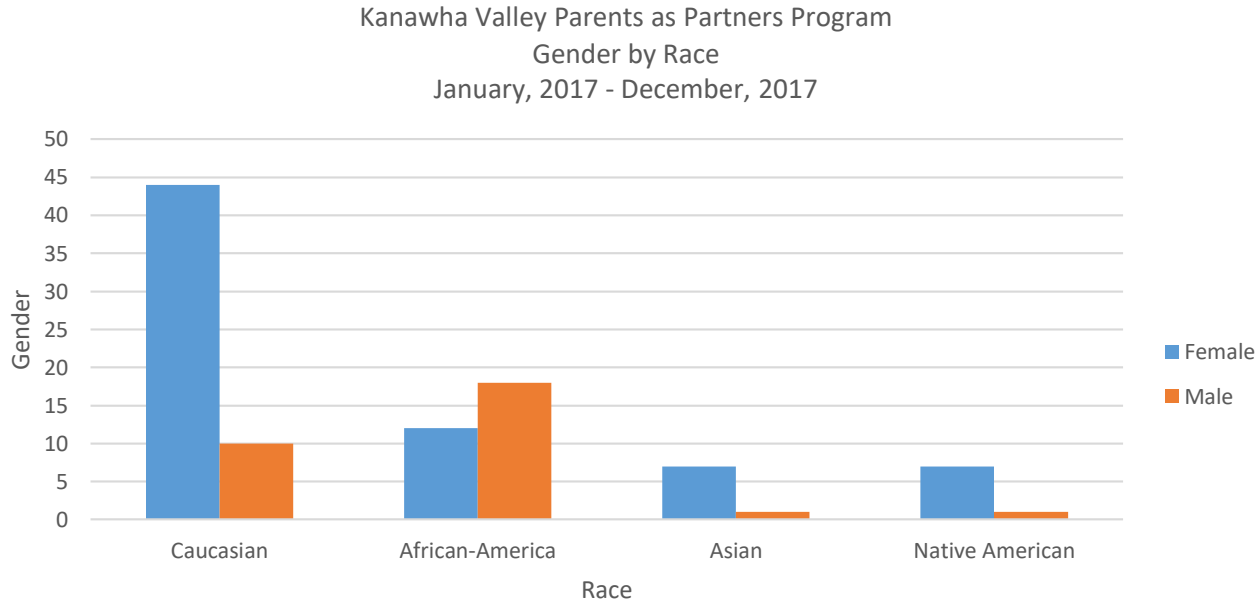
- Charts and graphs are used to portray:
 - Trends, relationships, and comparisons
- The most informative are simple and self-explanatory

Use The Right Type Of Graphic

- **Bar Chart:** Comparisons, categories of data
- **Line Graph:** Display trends over time
- **Pie Chart:** Show percentages or proportional share

Bar chart

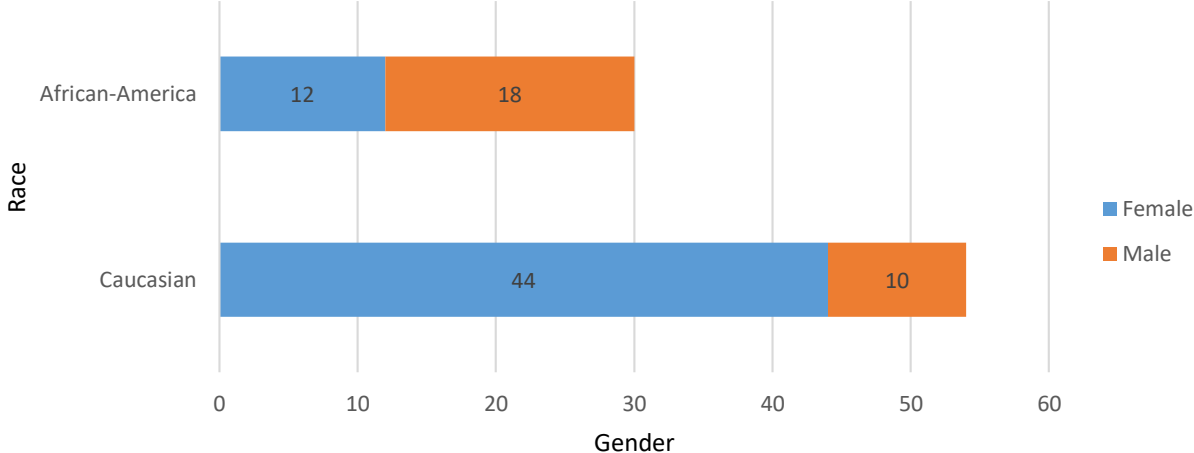
Comparing variables



Stacked bar chart

Represent components of whole & compare wholes

Kanawha Valley Parents as Partners Program
Number of Female and Male Participants by Race
January, 2017 - December, 2017

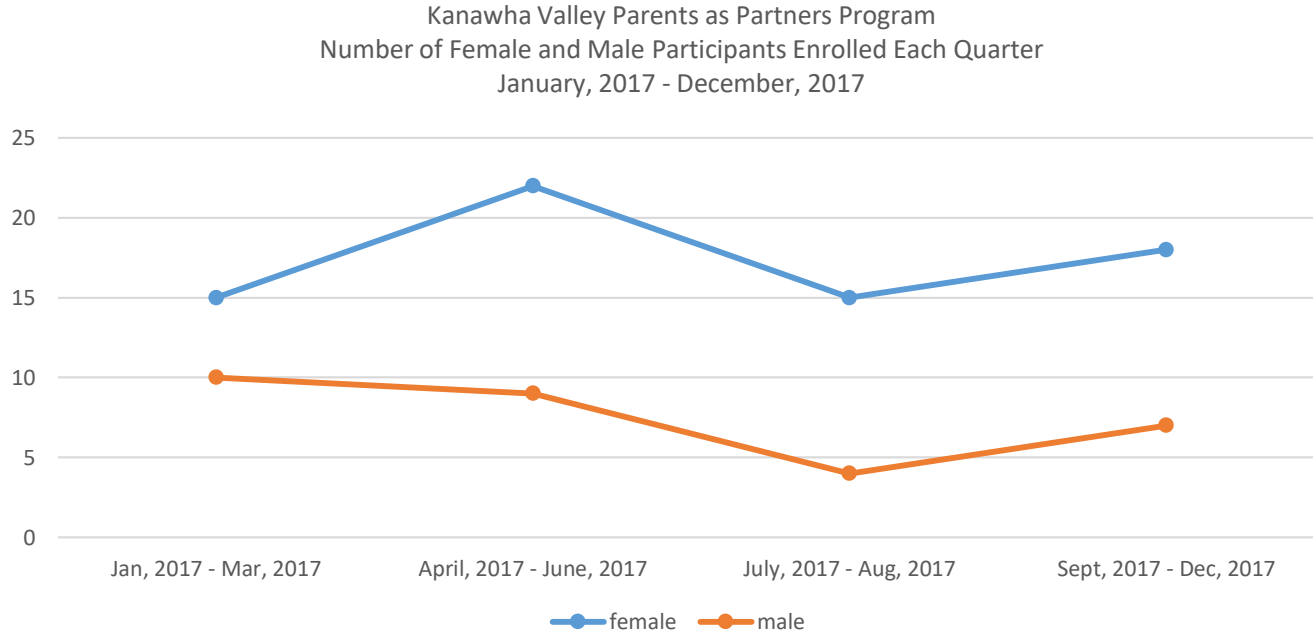


Midwest
Evaluation and
Research, LLC

Data Source: Kanawha Valley Parents as Partners Program Intake – January, 2017 – December, 2017

Line graph

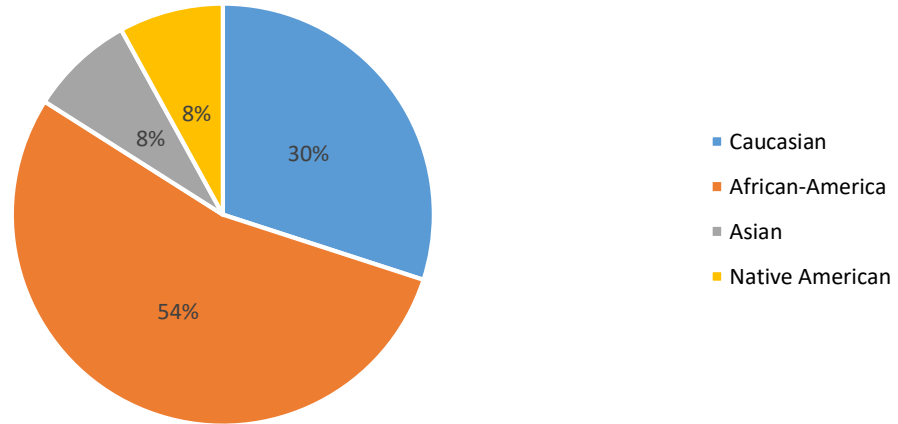
Displays trends over time



Pie chart

Contribution to the total = 100%

Kanawha Valley Parents as Partners Program
Percentage of All Participants by Location
January, 2017 - December, 2017




Questions



MER

Midwest
Evaluation and
Research, LLC



References

- *Basic Data Analysis (2012)*, Retrieved from <http://www.capacity4health.org>
- *Evaluation Toolkit (2017)*, Retrieved from <http://toolkit.pellinstitute.org>
- *Integrating Data Demand and Use into a Monitoring and Evaluation Training Course Training Tool Kit*, Retrieved from <https://www.measureevaluation.org>

Basic Excel Formulas

Excel Function	Formula	Example
Average	=average(data_range)	=average(B2:B20)
Sum	=sum(data_range)	=sum(B2:B20)
Mode	=mode(data_range)	=mode(B2:B20)
Median	=median(data_range)	=median(B2:B20)
Count	=count(data_range)	=count(B2:B20)
Countif	=countif(data_range, criteria)	=countif(B2:B20, 5)