MLE Measurement, Learning & Evaluation Project for the Urban Reproductive Health Initiative

Data Visualization Training Package

June 2017

MLE Measurement, Learning & Evaluation Project for the Urban Reproductive Health Initiative

Data Visualization Training Package:

The Measurement, Learning & Evaluation (MLE) Project for the Urban Reproductive Health Initiative

Authors:

Jack Hazerjian, Consultant, MLE Project

Christina Shaw, Johns Hopkins Center for Communication Programs

Rebecca Shore, Johns Hopkins Center for Communication Programs

Design:

Mark Beisser, Johns Hopkins Center for Communication Programs

Photo Credits:

2016, Tushar Sharma, Courtesy of Photoshare

2016, Linnea Zimmerman/PMA2020, Courtesy of Photoshare

2017, Courtesy of QuizFactor

2014, Courtesy of TechRepublic

2014, Courtesy of SAS 2010 Anil Gulati, Courtesy of Photoshare

Table of Contents

Acknowledgments 4

Data Visualization Training Package Overview 5

Training Agenda 6

Pre-Test 17

What is Data Visualization? 24

Types of Data Visualization 32

Data Visualization Checklists 44

Post-Test 49

Additional Resources 56

Appendix I 59

Acknowledgments

This report was made possible by support from the Bill & Melinda Gates Foundation under terms of the Measurement, Learning and Evaluation (MLE) Project for the Urban Reproductive Health Initiative. The MLE project is implemented by the Carolina Population Center at the University of North Carolina at Chapel Hill, in partnership with African Population and Health Research Center, International Center for Research on Women, Johns Hopkins Center for Communication Programs and Population Reference Bureau. The authors' views expressed in this publication do not necessarily reflect the views of the donor, the Bill & Melinda Gates Foundation.

What is the MLE Data Data Visualization Training Package?

In 2016, the Measurement, Learning & Evaluation (MLE) Project hosted two data visualization workshops for Bill & Melinda Gates Foundation project partners and stakeholders, as well as officials from Ministries of Health, in Abuja, Nigeria and Naivasha, Kenya.

The workshops were practical in-person trainings based on the Global Health eLearning Course, Data Visualization - An Introduction, created by former MLE Program Manager, Libby Skolnik and DHS colleague, Erica Nybro. The objective of the trainings was to enhance the capabilities of country partners to draw meaningful, practical information from data to ensure that health initiatives are evidence-based in their design and management.

This PDF contains the packaged version of the training resources from the two-day workshops, which includes a sample agenda, PowerPoint presentation, and training handouts, which are available in both English and French.

Training Agenda

The two-day data visualization training workshops were facilitated by MLE Consultant, Jack Hazerjian. The training agenda should be used as a guide for the facilitator to track the topics to be covered each day and to keep time to ensure all material will be covered.

TRAINING in DATA VISUALIZATION for
PROGRAM MANAGERS & RESEARCHERS in FAMILY PLANNING and REPRODUCTIVE HEALTH
DATE OF TRAINING **LOCATION OF TRAINING**

FIRST DAY
Date

Session Times	Session Topic	Session Objectives	Proposed Subject and Methodology	Training Aids
08:30 -- 09:00	**Pretest**		Before the start of the training, time is set aside for registration and seating taking; participants also take a Pretest (30 minutes)	➢ Pretest
09:00 – 09:05	**Welcome; Introductions**		Greetings and introduction of data visualization trainer to participants	
09:05 – 09:35	**Official Welcome; Opening Remarks**		Greetings and formal opening of training	
09:35 – 10:05	**Introduction**	**At the end of this session, participants will be able to:** • Define the term: "data visualization" • Describe the design of this training and the expected outcomes	Official welcome by trainer and introductions (10 minutes) Training facilitator ("Facilitator") welcomes the participants and provides an overview of the training program with a slide show on: • Training methods • Training objectives • Summary review of technical words, including the definition of the term, "data visualization" • Four procedural steps in creating a data visualization (30 minutes)	➢ Distribution of copies of "Training Agenda" ➢ PowerPoint Slides 02-14

Session Times	Session Topic	Session Objectives	Proposed Subject and Methodology	Training Aids
10:05 – 10:40	**Introduction** **(continued)**	• Rank participants' prior experiences in data visualization • Increase their experience through tasks assigned to them throughout the training	First Interactive Activity: Facilitator asks participants to evaluate their own level of confidence in creating data visualizations using the following classifications: • Are you crawling? • Are you walking? • Are you running? • Are you flying? Each participant presents herself/himself (name, work responsi-bilities) and reports on her/his level of confidence to do data visualizations on a Post-It; this piece of paper is put on a bar graph to show the variance among the responses of self-evaluation (10 minutes) Second Interactive Activity: Facilitator asks participants to separate into pairs in order to undertake a task: create a method for ranking in a graphical way the quality of instruction of each training session (morning and afternoon); each pair should document its evaluation of sessions in this way throughout the training (25 minutes)	➢ Post-Its and Markers ➢ Easel Paper, on which a bar graph depicts the first set of participant responses ➢ Paper ➢ Pen PowerPoint Slides 15 PowerPoint Slides 16
10:40 -- 10:55	**Coffee Break**			

Session Times	Session Topic	Session Objectives	Proposed Subject and Methodology	Training Aids
10:55 – 11:10	**Chapter 1: Identifying a Target Audience**	**At the end of this session, participants will be able to:** • List out the important characteristics to consider before developing a data visualization	**1st Step** – Facilitator presents the Matrix Chart of Target Audiences • Types of target audiences to be reached by those involved with health programming • Level of numeracy and data analysis skills of each of these different target audiences • Data sets that may be of particular interest to each of these different target audiences (15 minutes)	➢ Matrix Chart of Target Audiences ➢ Two sheets of Easel Paper, one titled: ○ types of target audiences ○ what one would like to know about target audiences ➢ PowerPoint Slides 17-21
11:10 – 13:00	**Chapter 1: Identifying a Target Audience**	**At the end of this session, participants will be able to:** • Respond to questions about case studies to demonstrate a good understanding of the context in which to prepare communications for a target audience	Third Interactive Activity: Facilitator leads a brainstorming discussion to determine: • Types of target audiences that the participants themselves may need to reach through their own work activities • Level of numeracy and data analysis skills of each of these different target audiences (20 minutes)	➢ PowerPoint Slides 23
			2nd Step – Facilitator presents the notion of context regarding the reception of data visualizations and asks the participants in plenary how the different target audiences could benefit from the data findings (20 minutes)	➢ PowerPoint Slides 24-33
			Fourth Interactive Activity: Facilitator asks each pair of participants, once regrouped, to study a data table example and identify possible target audiences and approaches to presenting findings from this table based on their numeracy and analysis skills (25 minutes)	➢ PowerPoint Slides 34-38
			Fifth Interactive Activity: Remaining in pairs, participants are asked to study one of two MLE Project data tables from Kenya and to identify possible target audiences and approaches to presenting findings based on their numeracy and analysis skills (45 minutes)	➢ PowerPoint Slides 39-42
13:00 – 14:00	**Lunch Break**			

Session Times	Session Topic	Session Objectives	Proposed Subject and Methodology	Training Aids
14:00 – 14:30	**Chapter 2: Identification of a Data Story to Tell to Others**	**At the end of this session, participants will be able to:** • Demonstrate an understanding of the terms: patterns, trends, comparisons, and outliers	**1st Step** – Facilitator provides a brief review of important terms -- patterns, trends, comparisons, ranges, outliers, surprises – by which to provide context for presenting data findings (15 minutes) First Interactive Activity: Facilitator solicits participants to study examples of data graphs and to identify which terms are most appropriate in reporting on the findings from them (15 minutes)	➢ PowerPoint Slides 45-54 ➢ PowerPoint Slides 55-69
14:30 – 14:55		• Demonstrate an understanding of the method of storyboarding	**2nd Step** – Facilitator provides a brief review of the process for creating a storyboard based on data (10 minutes) Second Interactive Activity: Facilitator displays the data table example and asks participants to work in their pairs to: • Write on Post-Its the different conclusions they can draw from this data graph • Rearrange the order of these Post-Its in order to interconnect their conclusions so that an overall summary statement can be made in the form of one sentence (15 minutes)	➢ PowerPoint Slides 70-74 ➢ Post-Its ➢ Pens and Markers ➢ Easel Paper ➢ Cellophane Tape ➢ PowerPoint Slides 75-76
14:55 – 15:10	**Coffee Break**			
15:10 – 16:30	**Chapter 2: Identification of a Data Story to Tell to Others (continued)**	• Facilitate a connection between a target audience and a data story	Third Interactive Activity: Facilitator displays seven MLE Project data tables from Kenya and asks participants to work in their pairs to select one or more data tables by which to put into practice the steps learned thus far during this training: • Select a particular, targeted audience and specify its level of numeracy and data analysis skills and its informational needs • Determine if the data table(s) reveals a pattern, a trend, a comparison, a range, outlying values, a surprise • Undertake the storyboarding process in order to tie together these conclusions to each other • Formulate an overall and definitive summary statement based on meaningful findings from your analysis of the data table(s) • Make a case to the participants in plenary for the effectiveness of this summary statement in reaching its target audience (80 minutes)	➢ Copies of Data Tables ➢ Post-Its ➢ Pens and Markers ➢ Cellophane Tape ➢ PowerPoint Slides 77-86

Session Times	Session Topic	Session Objectives	Proposed Subject and Methodology	Training Aids
16:30 -- 17:00	**Reflections on the Day**	1. Identify the key takeaways from today's sessions	Facilitator presents a summary of the main topics covered today (20 minutes) Interactive Activity: participants separate into pairs in order to rank in a graphical way the quality of instruction of today's training sessions (morning and afternoon) (10 minutes)	➢ Easel Paper, Markers ➢ Paper, Pen ➢ PowerPoint Slides 88

Session Times	Session Topic	Session Objectives	Proposed Subject and Methodology	Training Aids
09:00 -- 09:20	**Looking Behind and Ahead**	At the end of this session, participants will be able to: 1. Recall the important topics presented yesterday 2. Speak about topics to be presented today	Facilitator asks participants in plenary to review the topics presented yesterday and each pair of participants to share its evaluation of yesterday's morning and afternoon sessions Facilitator presents a summary of topics to be covered today (20 minutes)	➢ Easel Paper ➢ Markers ➢ Cellophane Tape ➢ PowerPoint Slide 2
09:20 -- 10:30	**Chapter 3: Development of a Data Visualization: Scientific Aspects on Perception, Applying what has been Learned**	At the end of this session, participants will be able to: 1. Demonstrate an under-standing of the four principles of data visualization, including: • Selection of data graphs • Simplicity in presentation style • Selective choice of pre-attentive attributes • Honesty	**1st Step** – Facilitator presents to the participants in plenary slides on: • Scientific aspects of visual perception (20 minutes)	➢ PowerPoint Slides 3-16
			First Interactive Activity: Facilitator asks participants to assess examples of data visualizations that **do not respect** scientific aspects of visual perception (10 minutes)	➢ Easel Paper & Markers ➢ Cellophane Tape ➢ PowerPoint Slides 17-20
			2nd Step – Facilitator presents to the participants in plenary slides on: • Four key principles to follow when developing a data visualization: ○ Selection of data graphs ○ Simplicity in presentation style ○ Selective choice of pre-attentive attributes ○ Honesty (40 minutes)	Copies of: ➢ "Selecting the Appropriate Data Graphic" ➢ "What Would You Like to Show" ➢ "Recommendations: Data Visualization" ➢ "Checklist for Assessing Data Visualization" ➢ PowerPoint Slides 21-67
10:30 -- 10:45	**Coffee Break**			
10:45 -- 12:00	**Chapter 3: Development of a Data Visualization: Applying what has been Learned (continued)**	At the end of this session, participants will be able to: 1. Apply the principles of perception and principles of data visualization in order to critique a data visualization	Second Interactive Activity: Facilitator leads a brainstorming discus-sion to critique and correct "bad examples" of data visualizations (20 minutes)	➢ PowerPoint Slides 69-81
			Third Interactive Activity: Facilitator asks each pair of participants to critique three data visualizations by identifying and listing out those aspects not following the four principles of good data visualizations and by recommending revisions to improve them (55 minutes)	➢ Easel Paper & Markers ➢ Cellophane Tape ➢ PowerPoint Slides 82-85

Session Times	Session Topic	Session Objectives	Proposed Subject and Methodology	Training Aids

12:00 – 13:00	**Chapter 3: Development of a Data Visualization, Applying what has been Learned (continued)**	At the end of this session, participants will be able to:
		1. Identify a data story for a selected target audience
		2. Develop an effective data visualization that conforms to four principles of data visualization and clearly transmits the intended message to target audience
			Fourth Interactive Activity: Facilitator engages the group in plenary to provide a brief review of all the topics presented over the course of the training up to the present for documentation on easel paper to serve as reference for the participants: (5 minutes)
			• Scientific aspects of visual perception
			• Four key principles to follow when developing a data visualization
			• Three chapters covered during the training sessions thus far
			  ○ Identification of a target audience
			  ○ Identification of a data story
			  ○ Development of a data visualization (15 minutes)
				➢ Easel Paper
				➢ Markers
				➢ Cellophane Tape
				➢ PowerPoint Slides 89-91

			Fifth Interactive Activity: Each pair of participants:
			• Receives an assigned target audience (e.g., policy maker, funder, journalist) with whom to engage with a data visualization
			• Selects from several, prepared data tables from which research-based information is to be shared with that target audience
			• Develops one or more data visualizations for that target audience
			• Documents on a easel paper how the data visualization(s) has/have been prepared, including:
			  ○ Data story and summary statement, as developed through storyboarding
			  ○ Scientific aspects of perception that are reflected in the design of the data visualization
			  ○ Four principles of data visualization that are likewise reflected
			• Exchanges its data visualization(s) with that/those of another pair who had been assigned a different role
			• Documents on easel paper the reasons it believes that the other pair must have had for developing its data visualization(s)
			• Posts side by side the two sets of easel paper offering explanations and reports results to the plenary group (40 minutes)
				➢ Easel Paper
				➢ Paper
				➢ Pen
				➢ Markers
				➢ Cellophane Tape
				➢ PowerPoint Slides 92-100

13:00 -- 14:00	**Lunch Break**			

Session Times	Session Topic	Session Objectives	Proposed Subject and Methodology	Training Aids
14:00 – 15:20	**Chapter 3: Development of a Data Visualization, Applying what has been Learned (continued)**	At the end of this session, participants will be able to: 3. Identify a data story for a selected target audience 4. Develop an effective data visualization that conforms to four principles of data visualization and clearly transmits the intended message to target audience	Fifth Interactive Activity (continued): Each pair of participants: • Receives an assigned target audience (e.g., policy maker, funder, journalist) with whom to engage with a data visualization • Selects from several, prepared data tables from which research-based information is to be shared with that target audience • Develops one or more data visualizations for that target audience • Documents on a easel paper how the data visualization(s) has/have been prepared, including: ○ Data story and summary statement, as developed through storyboarding ○ Scientific aspects of perception that are reflected in the design of the data visualization ○ Four principles of data visualization that are likewise reflected • Exchanges its data visualization(s) with that/those of another pair who had been assigned a different role • Documents on easel paper the reasons it believes that the other pair must have had for developing its data visualization(s) Posts side by side the two sets of easel paper offering explanations and reports results to the plenary group (80 minutes)	➢ Easel Paper ➢ Paper ➢ Pen ➢ Markers ➢ Cellophane Tape ➢ PowerPoint Slides 92-100
15:20 -- 15:35	**Coffee Break**			

Session Times	Session Topic	Session Objectives	Proposed Subject and Methodology	Training Aids
15:35 -- 16:05	**Chapter 3: Development of a Data Visualization, Applying what has been Learned (continued)**	At the end of this session, participants will be able to: 1. Identify the different types of data visualizations 2. Identify the different considerations in the design of each type of data visualization	**1st Step** – Facilitator presents to the participants in plenary some slides on: • Different types of data visualizations (except for those in reports) ○ Infographics ○ Conference posters ○ Dash boards ○ PowerPoint presentations ○ Web-based Videos ○ Film • Different considerations in the design of each type of data visualization (30 minutes) • ○ Infographics: logical flow in the presentation of several graphical elements, effective relationship among these elements regarding their size and placement, meaningful use of colors and fonts (size, style) ○ Conference posters: same considerations as above, plus those concerning scale of size of the poster itself ○ Dashboard: same considerations as those for Infographics, plus the issue of flexibility in data manipulation ○ PowerPoint presentation: logical flow in the presentation of the series of slides, effective relationship between text and images with regard to size and placement, consistent use of colors and font characteristics ○ Web-based video: logical flow and good pacing in the presentation of images in a series, effective relationship between text and images with regard to size and placement, meaningful use of colors, font, music and other sounds; plus the issue of flexibility in data manipulation ○ Film: same considerations as those for web-based video, except flexibility in data manipulation and scale of screen size	➢ PowerPoint Slides 102-116

Session Times	Session Topic	Session Objectives	Proposed Subject and Methodology	Training Aids
16:05 -- 16:40	**Chapter 4: Dissemination of a Type of Data Visualiza-to a Target Audience and Appraisal of the Results**	At the end of this session, participants will be able to: 1. Demonstrate an under-standing of the considera-tions when disseminating different types de data visualizations	Facilitator provides brief introduction to the final chapter (5 minutes) Facilitator presents slides to the participants in plenary on: • Considerations when disseminating any data visualization: ○ **Purpose** of the Data Visualization (keep in focus that targeted message is to be transmitted to elicit a desired reponse) ○ **Plan the Dissemination** (be specific in how to measure level of success in getting targeted message interpreted as desired and response elicited as intended) ○ **Perception & People** (don't forget that the various audience groups interpret messages differently and respond to them differently ○ **Process** of the Data Visualization Development (ensure that all processes are monitored: data quality, diversity among developers of data visualization, level of scrutiny in discerning and selecting key messages and in critiquing design of data visualization (30 minutes)	➢ PowerPoint Slides 117-128
16:40 -- 17:10	**Reflections on the Day**	At the end of this session, participants will be able to: 1. Identify the key takeaways from today's sessions	Facilitator presents a summary of the main topics covered today and reference to resources to learn more about data visualization (20 minutes) Interactive Activity: participants separate into pairs in order to rank in a graphical way the quality of instruction of today's training sessions (morning and afternoon) (10 minutes)	➢ Easel Paper ➢ Markers ➢ Paper, Pen ➢ PowerPoint Slides 129-133

Pre-Test

The pre-test should be given at the beginning of the first day of the workshop. The purpose is to test participants on their knowledge of data visualization approaches and best practices. Following the workshop, the participants will take this again as a post-test to see the knowledge they gained during the two-day training

MLE PROJECT: **TRAINING in DATA VISUALIZATION**

PRE-TEST

Please check off your response for each of the following questions
– unless indicated differently by the instructions

1. Data visualization is a way of rebalancing so that information transmission is promoted through visual perception, taking advantage of the powerful capacity of eyes to scan information.

☐ True

☐ False

2. Data visualization is primarily concerned with graphic design.

☐ True

☐ False

3. Charts and graphs developed in the usual way are not considered data visualizations.

☐ True

☐ False

4. The principal point to make (related to findings from a data analysis) should not be directly specified in a data visualization.

☐ True

☐ False

5. From a set of data, there is only one principal point or finding that can be shared with a target audience that can be recognized as a correct and meaningful interpretation.

☐ True

☐ False

PRE-TEST (continued)

6. Data visualizations greatly serve to disseminate the sense of the data ("what does it mean?") and to encourage data use for corrective action.

- [ ] True
- [ ] False

7. When presenting a percentage distribution chart, a pie chart is always recommended to be used.

- [ ] True
- [ ] False

8. In a graphical presentation of the data, it is commonly recommended to create a chart in three-dimensional form in order to reinforce its visual impact.

- [ ] True
- [ ] False

9. In a graphical presentation created in several colors, one should consider the potential problem posed by their different combinations for those who are colorblind.

- [ ] True
- [ ] False

10. In a graphical presentation created in several colors, one does not need to consider the cultural associations of colors.

- [ ] True
- [ ] False

PRE-TEST (continued)

11. Using the numbers «1» to «4», put in chronological order the following four steps concerning the design process for a data visualization.

- [ ] Analyze the data to determine if there can be found a pattern, a trend, a comparison, or outlier values; from these observations, seek out a meaningful interpretation of the data
- [ ] Sketch out with pencil and paper several ideas that come to you to assess the best approach for transmitting the data analysis findings with a data visualization, recognizing that visual approaches often make information more understandable and compelling than explanatory text
- [ ] Determine your target audiences and stakeholders with whom the data visualization will be shared and identify what each of them should of them should understand from the data and data finding
- [ ] Before the dissemination of a data visualization, customize it for each target and for each method of presentation (report charts are done differently than images for an infographic or a video), recognizing that the necessary data needs to be updated

12. Indicate with a check mark if each of the following items is **True** regarding who are the different target audiences for a data visualization about program data.
[There may be more than one correct response among the four items.]

- [ ] Program managers
- [ ] Program service providers
- [ ] Journalists
- [ ] Researchers

13. Indicate with a check mark if each of the following items is **True** regarding the considerations to keep in mind for each of the different target audiences when developing a data visualization about program data.
[There may be more than one correct response among the four items.]

- [ ] Average level of skills in numeracy (math calculations)
- [ ] Average level of skills in program design
- [ ] Average level of skills in graphic design
- [ ] Average level of skills in data interpretation (data analysis)

PRE-TEST (continued)

14. Indicate with a check mark if each of the following items is **True** regarding a data visualization for a program.
[There may be more than one correct response among the four items.]

- [ ] It is a presentation of complex ideas in the form of a story
- [ ] It is a presentation of findings with proper consideration for graphic display
- [ ] It is a presentation of information done in a way to facilitate comprehension
- [ ] It is a presentation of data done to engage with a certain target audience and to instill a response for corrective action

15. Indicate with a check mark if each of the following items is **True** regarding the means for enhancing transmission of meaningful findings by a data visualization.
[There may be more than one correct response among the four items.]

- [ ] Colors are carefully selected
- [ ] All text font is boldfaced
- [ ] Chart legend is placed to the side in order to free up space in the middle
- [ ] Chart title summarizes data findings and is written to highlight their significance

16. Indicate with a check mark if each of the following items is **True** regarding the development of a data visualization.
[There may be more than one correct response among the four items.]

- [ ] Chart axis showing values of magnitude should begin at zero
- [ ] Tick marks should be removed from axes, unless they are necessary
- [ ] Grid lines should be clearly visible in a chart
- [ ] Label and legends should have borders to better encapsulate their words

PRE-TEST (continued)

17. Indicate with a check mark if each of the following items is **True** regarding the development of an infographic. *[There may be more than one correct response among the four items.]*

- [ ] Series of images used in an infographic should run from left to right
- [ ] Expected understanding of an infographic by the target audience should be considered
- [ ] Infographic should be developed by a team of analysts and graphic designers
- [ ] Interpretation to be transmitted by an infographic should be first formulated by means of storyboarding

18. Match each term with its definition found to the right.
[Document your choices by writing the alphabet code in the box in front of each term.]

[ ]	Pattern	**A**	Detection of similarities and differences among data points in a data set
[ ]	Trend	**B**	Detection of the minimum and maximum values in a data set
[ ]	Comparison	**C**	Detection of a regular and repeated relationship in the distribution of data points in a data set
[ ]	Range	**D**	Detection of a distribution of data points that conforms to a certain arrangement over the course of time

19. Match each chart type with the reference found to the right concerning the type of interpretation of data to be presented. *[Document your choices by writing the alphabet code in the box in front of each chart type.]*

[ ]	Scatter Plot	**A**	Comparison of total fertility rates among five countries
[ ]	Line Graph	**B**	Distribution by age in a country over the course of time
[ ]	Pie Chart	**C**	Distribution by sex where 60% of the national population is female
[ ]	Bar Chart	**D**	Total fertility rate with respect to prevalent utilization rate of contraceptives in each of twenty countries
[ ]	100% Area Chart	**E**	Fertility rate trend over the course of time in each of several countries

PRE-TEST (continued)

20. Match each type of technical specialization with the reference found to the right of tasks involved in the development of a data visualization.
[Document your choices by writing the alphabet code in the box in front of each type of specialization.]

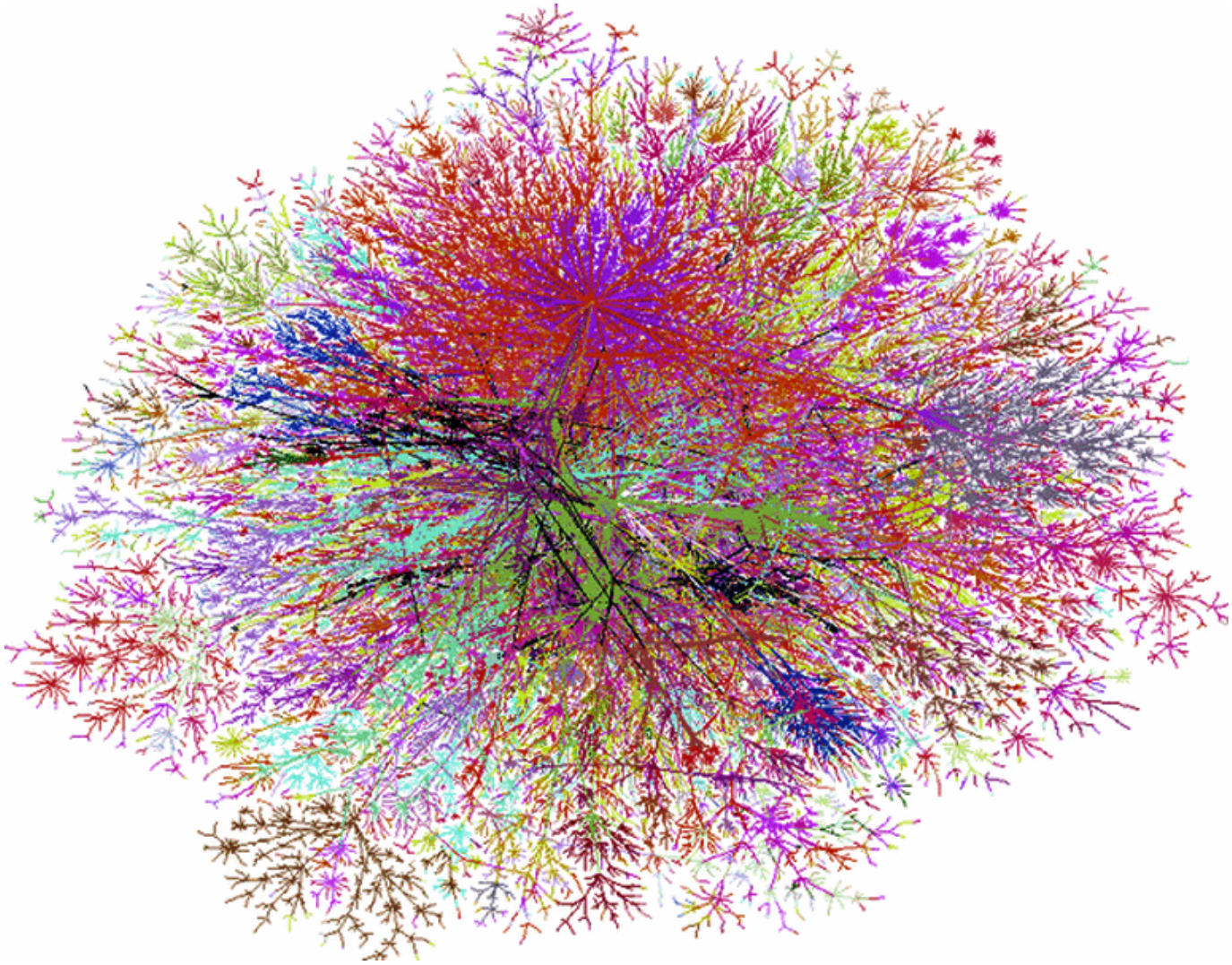
☐	Data Quality Monitoring	**A**	Presentation of an interpretation of the data in a way that is appropriate for the target audience
☐	Technical Programming	**B**	Collection of accurate and complete data for the visualization and quality control of the data interpretation in the visualization itself
☐	Communications	**C**	Quality control on the reporting of programming carried out in relation to the visualization's interpretation of the data that is being transmitted
☐	Computer Programming	**D**	Identification and refinement of messages to be remembered by target audiences from the data visualization
☐	Graphic Arts	**E**	Adaptation of a data visualization so that it can be transmitted on another platform (for example, on the Web)

21. Match each type of data visualization with its definition found to the right.
[Document your choices by writing the alphabet code in the box in front of each term.]

☐	Infographics	**A**	Means for presenting dynamically and specifically quantitative data on demand with customized software
☐	Charts and Graphs	**B**	Means for broadly transmitting an interpretation of quantitative and qualitative data through a series of static images
☐	Videos, Multimedia	**C**	Means for presenting easily and precisely an interpretation of the data in geometric and static forms
☐	Interactive Web Sites	**D**	Means for dynamically transmitting an interpretation of the data through animated images and the use of narrative and sound

What is Data Visualization?

The next several resources explain what data visualization is and why it is important.



What is Data Visualization?

- **Varied definitions (generally, use of graphic means to translate data into a compelling story/meaningful interpretation of data results)**
- **Practice that recognizes**
  - Interplay of perception and cognition
    - Perception = unconscious and split-second process for understanding of what one sees
    - Cognition = conscious and deliberate (slower) process for understanding of what one sees
  - Limits of short-term memory
  - Benefits of pre-attentive attributes
  - Visual processing described in Gestalt theory principles
- **Practice that combines skills in three areas:**
  - Data analysis
    - Identifying and assessing key findings from the data to share
  - Communications
    - Identifying the different audiences with whom to share key findings
    - Identifying (possibly different) key findings to share with each of these audiences
    - Assessing the different levels of responsiveness to data visualizations by each of these audiences
  - Graphic design
    - Identifying appropriate methods for representing key findings
    - Developing clear and compelling visualizations of the data

Why is Data Visualization Important?

- **Promotes greater likelihood of data use for evidence-based decision making if data is presented in attention-grabbing manner**
- **Promotes an emphasis on the highlighting of key findings from the data, rather than a broad presentation of all the data**
- **Promotes the careful consideration of how data presentations will be viewed and understood by each of the different audiences**
- **Provides balances to information and design considerations**
  - Integrity
    - Accuracy and timeliness of information
  - Function
    - Relevance and usefulness of information
  - Interestingness
    - Attractiveness and compellingness of the design
  - Form
    - Clarity, legibility, and purposefulness of design

How to do Data Visualization?

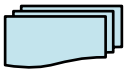
- **Identify main audience**
  -  Higher-level decision makers who oversee funding and support for services documented in the data
  -  Managers and administrators of these services
  -  Front-line staff who provide these services
  -  General public served by these services

-  Broader community of professionals and organizations involved with design, funding, implementation, and results of these services

- **Identify areas of interest / concern of main audience**
  -  Higher-level decision makers
    - Volume of services at national and regional levels (magnitude, change); coverage rate of services (national and regional populations); level of achievement to targets set for services at national and regional levels; knowledge about, attitudes towards, and usage of services (by national and regional populations)
    - National and regional levels of staffing (magnitude, position types, geographic and urban/rural presence, vacancies, training, credentials)
    - Expenditures for provision of services throughout the country and regionally (item costs, costs per unit, efficiency)
    - Status change in national and regional populations targeted for services due to service provision
  -  Managers and administrators
    - Volume of services under their responsibility (magnitude, change); coverage rate of services in their catchment areas (population and subpopulations therein); level of achievement to targeted levels of services in their catchment areas; knowledge about, attitudes towards, and usage of services in their catchment areas (population and subpopulations therein)
    - Staffing under their authority (magnitude, position types, geographic and urban/rural presence, vacancies, training, credentials)
    - Expenditures for provision of services in their catchment areas (item costs, costs per unit, efficiency)
    - Status change in population targeted for services due to service provision in their catchment areas
  -  Front-line staff
    - Volume of services staff has provided (magnitude, change); coverage rate of services in their catchment areas (population and subpopulations therein); level of achievement to targeted levels of services in their catchment areas; knowledge about, attitudes towards, and usage of services in their catchment areas (population and subpopulations therein)
    - Status change in population targeted for services due to service provision in their catchment areas
  -  General public
    - Volume of services provided nationally/regionally (magnitude, change); coverage rate of services (national and regional populations); level of achievement to targeted levels of services; knowledge about, attitudes towards, and usage of services (national and regional populations)
    - Status change in national and regional populations targeted for service provision
  -  Broader community of professionals and organizations
    - Volume of services (change); coverage rate of services (national and regional populations); level of achievement to targeted levels of services; knowledge about, attitudes towards, and usage of services (national and regional populations)
    - Expenditures for provision of services (efficiency)
    - Status change in population targeted for services due to service provision

- **Identify competence of main audience in reviewing data and interpreting data**
  -  Higher-level decision makers
    - Should have competency to broadly understand percentage changes related to provision, coverage, and usage of services as well as percentage changes in status of targeted population(s) — although limitations in abilities and time may require that conclusions drawn from data analysis be pre-formulated and given
    - Should have competency to broadly understand staffing-related issues, but may require that conclusions drawn from data analysis be pre-formulated and given
    - Should have competency to broadly understand service expenditures but may require that conclusions drawn from financial analysis be pre-formulated and given
  -  Managers and administrators
    - Should have competency to understand percentage changes related to provision, coverage, and usage of services as well as percentage changes in status of targeted population(s) within their catchment areas — and should conduct their own data analysis and draw their own conclusions for presentation to their superiors/subordinates
    - Should have competency to understand staffing-related data within their catchment areas — and should conduct their own data analysis and draw their own conclusions for presentation to their superiors/subordinates
    - Should have competency to understand service expenditures within their catchment areas— and should conduct their own financial analysis and draw their own conclusions for presentation to their superiors/subordinates
  -  Front-line staff
    - Should have basic competency to understand percentage changes related to provision, coverage, and usage of services as well as percentage changes in status of targeted population(s) within their catchment areas — although limitations in abilities and time require that conclusions drawn from data analysis be given to them so that responsive action related to service provision can be taken
  -  General public
    - Should have basic competency to understand percentage changes related to provision, coverage, and usage of services as well as percentage changes in status of targeted population(s) — although conclusions drawn from data analysis should be pre-formulated and given
  -  Broader community of professionals and organizations
    - Should have strong abilities to analyze and draw conclusions related to data analysis on provision, coverage, and usage of services as well as on status change of targeted populations — although limitations in time may require that conclusions be pre-formulated and given
- **Find story in data of interest to main audience**
  -  Higher-level decision makers = positive/negative changes at national and regional levels in:
    - Services (volume, coverage, meeting targeted levels, utilization)
    - Staffing (credentialing, hiring, placing, retaining)
    - Expenditures for services (spending levels, cost savings, efficiencies)
    - Status change of populations targeted for services

-  Managers and administrators = positive/negative changes within their catchment areas in:
  - Services (volume, coverage, meeting targeted levels, utilization)
  - Staffing (credentialing, hiring, placing, retaining)
  - Expenditures for services (spending levels, cost savings, efficiencies)
  - Status change of populations targeted for services
-  Front-line staff = positive/negative changes within their catchment areas in:
  - Services (volume, coverage, meeting targeted levels, utilization)
  - Status change of populations targeted for services
-  General public = positive/negative changes at national and regional levels in:
  - Services (volume, coverage, meeting targeted levels, utilization)
  - Status change of populations targeted for services

- **Test out ways to tell that story through data visualization**
  -  Determine if data values present trends, correlations, and/or outlier
  -  Determine if data values are nominal (non-quantified labelling values)
    - Data values should be emphasized by position, shape, and coloration in the data visualization
  - Determine if data values are quantitative (numeric values)
    - Data values should be emphasized by position, length, angle, and area in the data visualization
  -  Determine if data values are ordinal (ranking values)
    - Data values should be emphasized by position, coloration, and length in the data visualization

- **Build data visualization to tell story**
  - Graphs
    - Time series: data values show change over time
    - Ranking: data values show order of magnitude, in ascending or descending format
    - Part-to-whole: data values show proportional relationships of parts to whole
    - Deviation: two sets of data values are cross-compared for statistical difference (i.e., variance)
    - Distribution: data values within each comparable subgroup cross-compared
    - Correlation: data values in the form of two variables are compared to determine the strength of their intra-relationship
    - Geospatial: data values are ranked and mapped in order to illustrate possible, geographic patterns
    - Nominal comparison: data values are cross-compared without a quantitative means of ranking
  - Dashboards
  - Infographics
  - Word clouds
  - Storyboards
  - Videos/animations
  - Interactive graphic websites

Good and Bad Practices in Data Visualization: Summary

- **General principles of design for data visualizations**
  - Use only one color to graphically represent a data variable category
  - Use a clearly understandable method for differentially ranking the data sets within a graph
  - Use the graph's title or text boxes within the graph in order to highlight the most significant finding(s) to share with the viewer
  - Ensure that the design and contents of a data visualization lend themselves for clear and immediate understanding of the findings
  - Use well-known icons and other graphical symbols whenever possible to reduce textual labels
  - Avoid the use of color combinations that are difficult to differentiate for those who are color blind and for those who are distracted by the color contrast:
    - green and red green and blue green and gray green and brown green and black light green and yellow
    - blue and yellow blue and purple blue and gray
  - Avoid the use of three-dimensional versions of charts, since they often lead to data misinterpretations
  - Avoid all unnecessary elements in data visualizations that serve merely as pictorial enhancements (e.g., arrows, frames, shadows)
  - Avoid the use of more than six colors in any one data visualization, since multiple colors can be visually distracting
  - Avoid the use of underlined, italicized, or boldfaced text

- **In short, keep it clear, concise, calm, uncluttered and compelling**

Good and Bad Practices in Data Visualization: Data Visualization Types

- **Area charts**
  - Use area charts to graphically represent a time-series (change over time) relationship of continuous data[1] in order to highlight the magnitude of differences among several sets of data points
    - Use stacked area charts to highlight the comparative contribution of each part to the cumulative whole
    - Use 100% stacked area charts to highlight the proportional contribution of each part to the cumulative whole
      - Begin the y-axis at zero, whenever possible
      - Stack the area sections from bottom to top in ascending order of the variability of the sections' respective "y" values; that is, the section with the most irregular contour line should be the topmost section, while the section with the flattest contour line should be on the bottom
      - Avoid the use of area charts for graphing discrete data[2]
      - Avoid the use of solid colors for area sections; use transparent color tones to allow for the overlapping of area sections to be visible
      - Avoid the use of more than four sections in an area chart

- **Bar charts**
  - Use bar charts to graphically represent continuous or discrete data by means of a time-series relationship, or a simple cross-comparison between different sets of data points, or a comparison of parts to the whole within each set of data points
    - Use stacked bars to graphically represent the relationship of multiple parts to the whole for sets of data points that can be discrete or continuous
    - Use 100% stacked bars to highlight the proportional distribution of each of part to the cumulative whole

---

[1] **Continuous data**: data points presented in numeric form that are not a discrete, defined set of values but ones that lie within a wide range possible values, such as degrees of temperature, measurements of height and weight, time durations

[2] **Discrete data**: data points presented in numeric or nominal form for each of which has a specified, distinct value with no intermediary

- Begin the y-axis at zero
- Scale the width of the space between bars to be one-half the width of a bar
- Maintain same color for bars except for the bar(s) to be differentiated from the others
- Use vertical bars to graphically represent a time-series relationship, with time advancing from left to right
- Position horizontal bars using a logical method (e.g., by bar length, by alphabetical order of bar label)

- Avoid the use of horizontal bars, unless the label for each bar would have been too long to use for bars oriented vertically
- Avoid the use of vertical labels for all bars, whether horizontally- or vertically-oriented; position labels within or close by those bars that run horizontally

- **Line charts**
  - Use line charts to graphically represent a time-series relationship of continuous data in order to highlight trends, levels of increase and decrease, and patterns of volatility

    - Begin the y-axis at zero, whenever possible, unless the values on the y-axis are too large to fit in the graph
    - Create a close up of a line chart when the difference in the "y" values are comparatively small to each other but nonetheless meaningful overall to show
    - Scale the length of the y-axis so that the largest "y" value of a data point is located about two-thirds of the way up the height of the line graph

    - Avoid the use of legends to indicate what each line denotes; label the lines directly in the graph itself
    - Avoid the use of lines made with dots or dashes; solid lines are best
    - Avoid the use of more than four lines in a line chart

- **Pie charts/donut charts**
  - Use pie charts/donut charts to graphically represent the relationship of multiple parts to the whole for a <u>small</u> (up to five) sets of data points that can be discrete or continuous

    - Percentage values of pie/donut wedges should add up to 100%
    - Position pie/donut wedges in either of two ways:
      - Largest two wedges placed on either side of 12 o'clock position, with largest wedge running clockwise while the second largest and the remaining wedges run counterclockwise by order of descending size
      - Largest wedge placed at 12 o'clock position and running clockwise, followed by the remaining wedges running in the same clockwise direction by order by descending size

    - Avoid the use of more than five wedges in a pie/donut chart
    - Avoid the use of multiple pie/donut charts for cross comparison

- **Scatter plot charts**
  - Use scatter plot charts to graphically represent the degree of correlation among a large number of data points each of which, depicted as a dot, is defined by a pair of variables whose values can be discrete and/or continuous

    - Use simple scatter plot charts when both variables have continuous values
    - Use scatter plot jittering (with random noise added to the data) or box and whiskers when the values of only one of the variables are continuous (or have been converted to continuous from discrete)

    - Begin the y-axis at zero
    - Use a trend line to demarcate slope of bivariate correlation

    - Avoid the use of more than two trend lines

- **Bubble charts**
  - Use bubble charts to graphically represent comparative relationships among three variables whose values can be discrete and/or continuous or to graphically rank their relationships, especially in context of geographic location
    - Use varying size and colors to differentiate dots of a scatter plot chart when a third variables has been introduced
    - Ensure bubbles are clearly labelled
    - Ensure bubbles are proportionately sized according to their area --- and not their diameter
    - Avoid the use of non-circular (irregularly-shaped) bubbles
    - Avoid the use of bubbles containing more than the basic graphical information of size, color, label
- **Heat maps/conditional density plots**
  - Use heat maps/conditional density plots to graphically represent frequency occurrence of categorical data through the use of color gradations, especially in context of geographic location
    - Use heat maps/conditional density plots for mapping frequency of bivariate data points when both variables have discrete values
    - Use light-to-dark tonalities of the same color in order to indicate the category ranges of frequency occurrence of data points
    - Use only three to five categories by which to differentiate the category ranges of frequency occurrence
    - Use minimal visual means (thin boundary lines in a neutral color) for differentiating geographic locations
    - Avoid the use of more than pattern scheme by which to further differentiate the tonal color range of frequency occurrence

- **Disseminate and use data visualization**

- **Assess data visualization**
  - Evaluate occurrence of any of three possible types of communication problems (Shannon and Weaver)
    - Technical problem = less than full accuracy in presentation of data values and results by data visualization
    - Semantic problem = apparent difficulty of target audience in reading data values and drawing intended conclusions from data visualization
    - Effectiveness problem = failure of data visualization to cause target audience to meaningfully respond to data visualization and its intended conclusions

- **Adapt / revise data visualization for a different audience interested in the data**

Types of Data Visualization Graphics

The next few resources walk you through the different types of data visualization and when to use each one.

RECOMMENDATIONS FOR GOOD DATA GRAPHICS

TEXT

Data graphics should not contain too much text. Text of a data graphic should powerfully project a riveting message.

Efficacy of a Daily Dosage to Prevent HIV: Young African Women Regularly Taking Truvada Remain Non -Infected

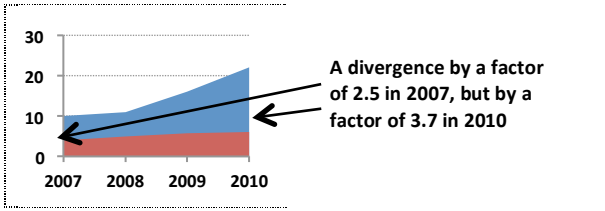
In a study conducted in three African countries since the start

Efficacy of a Daily Dosage to Prevent HIV: Young African Women Regularly Taking Truvada Remain Non -Infected

In a study conducted in three African countries since the start

Provide a descriptive title of 6-12 words and position it in the upper left hand corner with left justificati

A brief title allows the reader to comprehend and retain the principle message principal, even when skimming a data graphic. Instead of words with non-specificity, use a very descriptive phrase to capture the findings derived from a data graphic in order to highlight its importance. Since we read from left to right, position the title to begin as well from the left side



A subtitle or a comments box can serve to provide supplemental information

A subtitle or commentary (for example, a dialogue box) can enhance the effectiveness of a data graphic to explain or interpret the data. Use either one to respond to possible questions that a reader may have or to highlight a conclusion that one could draw from the data

Title
Subtitle
Commentary
Text/Data Label
Axis Label
Footnote

9 - 14
Font sizes used for paper reports
20 – 48
Font sizes used for projector slides

Font size is hierarchical and legible

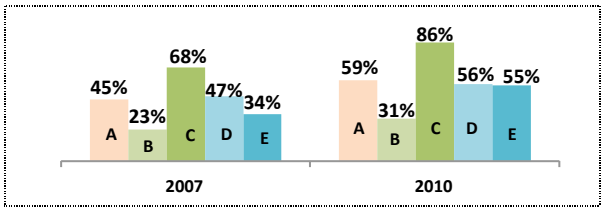
The title should be in a larger font size than that for a subtitle or dialogue box, which should be in a larger font size than that for text/data label; in turn, a text/data label should be in a larger font size than that for an axis label, while the smallest font size is reserved for footnotes. For reference, the axis label should have a font size of at least 9 points when printed on paper and 20 points when presented on a screen projection

Title
Subtitle
Commentary
Text/Data Label
Footnote

Axis Label

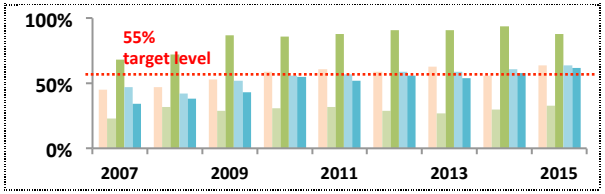
Text runs horizontally

Titles, subtitles, commentary, and text/data labels should always run horizontally (that is, neither vertically nor diagonally). On the other hand, one can ignore this rule for axis labels



Labelling data points

Position the label directly on elements representing data points or right next to the data points, instead of having a legend box some distance away (That is, place labels directly above the bars in a bar graph, outside the wedges of a pie chart, beside the lines in a line chart). Either eliminate altogether labels or integrate them in the data graphic because any distance between labels and data points can cause inattention and provoke difficulty in data interpretation



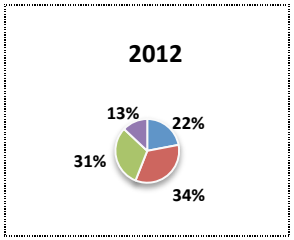
Use labels wisely

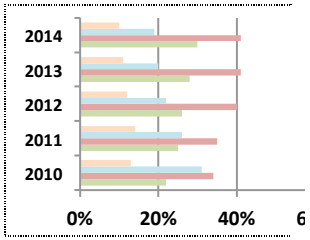
Facilitate attention being paid on the essentials of a data graphic by removing un-useful labels. For example, simply note every second year on an axis indicating time period for a data graphic

RECOMMENDATIONS FOR GOOD DATA GRAPHICS (2)

ARRANGEMENT

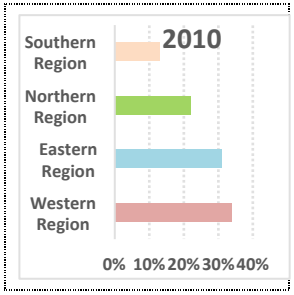
Erroneous arrangement of graphical elements can at best confuse readers and at worst create errors of comprehension for them.
A carefully laid out arrangement more easily allows for data interpretation.

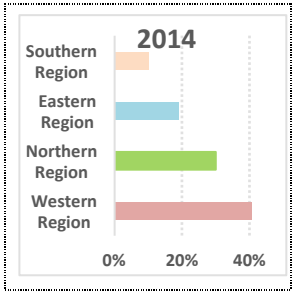




Proportions are accurate

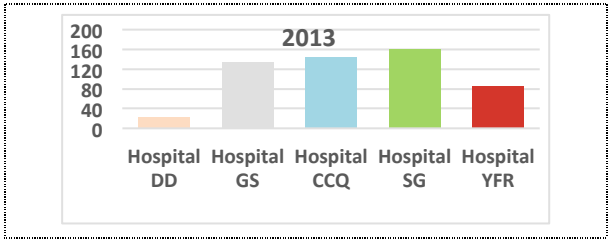
A person ought to be able to confirm with a ruler that proportional relationship among lines or areas represented in a data graphic do indeed correspond to the values indicated by their respective labels





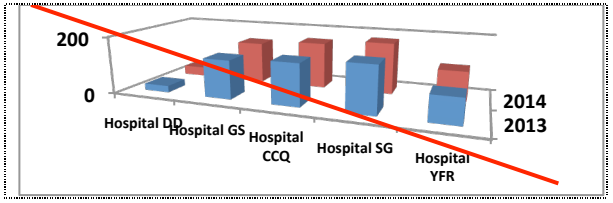
Data is presented in an intentional order

Data should be arranged in an order that is reasonable and logical. One could present data by means of a frequency count (that is, from largest to smallest for nominal categories), in alphabetical order, by groups or intervals (e.g., histograms), or with respect to the flow of time (line graph)



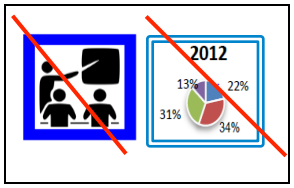
Interval spacings on an axis are always the same width

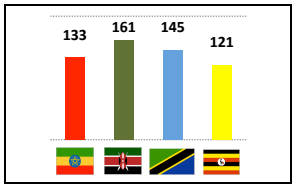
There is always the same distance between intervals on an axis, regardless whether the axis is labeled or not



Data graphic is always represented in two-dimensional form

Avoid using three-dimensional representations and other images with beveled edges and visual distortions





Data graphic is free from ornamentation

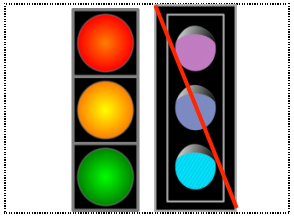
Avoid using Cliparts and other images (like borders) to merely decorate a data graphic.

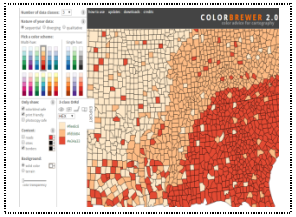
Nonetheless, it could be useful to add certain types of images, like icons, to reinforce the interpretation of data in a graphic

RECOMMENDATIONS FOR GOOD DATA GRAPHICS (3)

COLORS

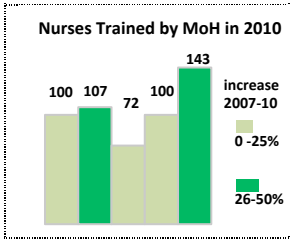
Consider cultural associations that colors may have (e.g., death can be linked to the color black or white.) Consult web sites like "Color Brewer" in order to identify color schemes that are recommended for photocopied images and for readability by those who are colorblind.

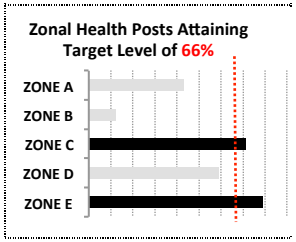




Color schemes arise from rational design considerations

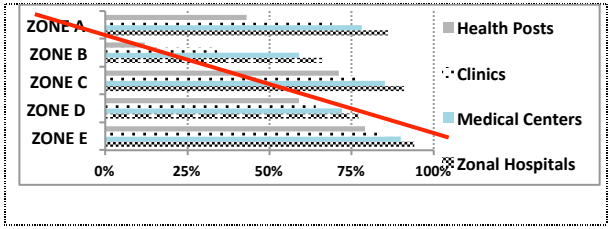
Colors are intentionally selected for specific reasons – e.g., they are associated to a brand logo related to the study subject represented by the data. Avoid relying upon a default color scheme, but then again make use of color schemes related to your subject. Moreover, one should do research on tools found online in order to identify colors associated with a brand and to select colors that are compatible with each other





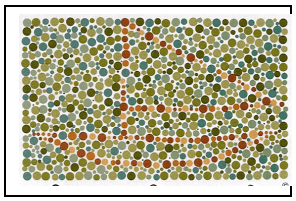
Colors help to highlight trends brought out by the data

Bright colors should be deployed to guide the viewing of a data graphic for the key results coming from the data. Use muted colors for elements in a data graphic that are less important or informative



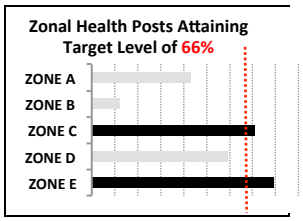
Perception of differences among colors after photocopying in black and white

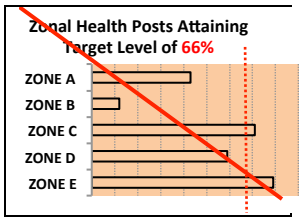
When printing or photocopying in black and white a data graphic that had been created in color, it is necessary that there remain differences of tonality among the colors in order to highlight the trends brought out by the data



Perception of differences among colors for the colorblind

Avoid placing side by side the colors red and green as well as yellow and blue since someone who is colorblind cannot easily distinguish one from the other





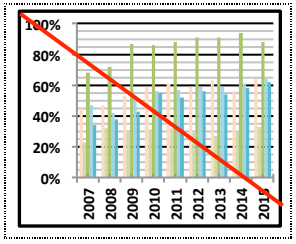
Text color in sufficient contrast with the background color

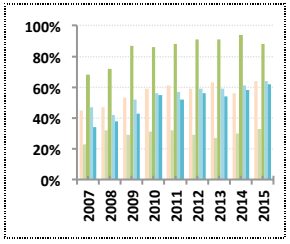
Viewing a data graphic is facilitated when the text color is black or a very dark color while the background color is white or transparent

RECOMMENDATIONS FOR GOOD DATA GRAPHICS (4)

LINES

As a group, grid lines, borders, tick marks, and axes can all create visual disorder in a data graphic. Accordingly, minimize their inclusion if they do not serve to facilitate the interpretation of the data.

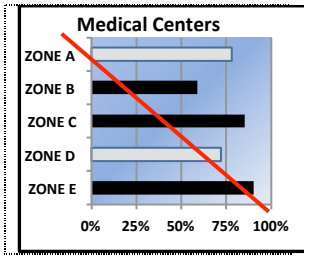


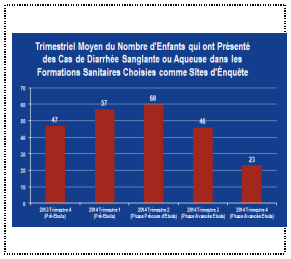


Grid lines in muted a color, if they are retained in a data graphic

Use light gray, not black as a color for the grid lines.

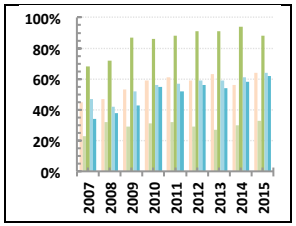
Better still, do not use grid lines at all

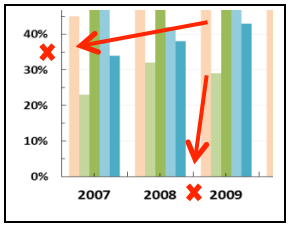




No borders for a data graphic

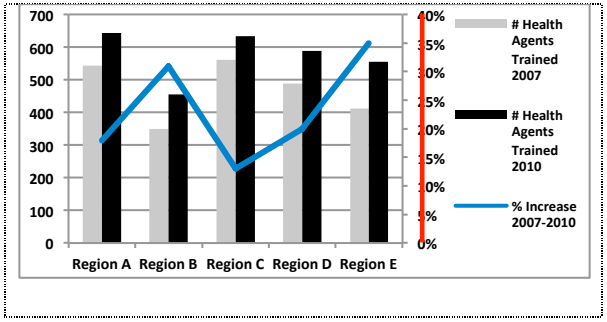
The background of a data graphic should merge invisibly with the surrounding area on the page or slide instead of being framed by a border





Lack of tick marks on the axes, unless they are necessary

Tick marks are useful for line graphs (to mark off each point in time on the vertical axis); however, tick marks are not useful for bar charts



Restriction to one horizontal and vertical axis

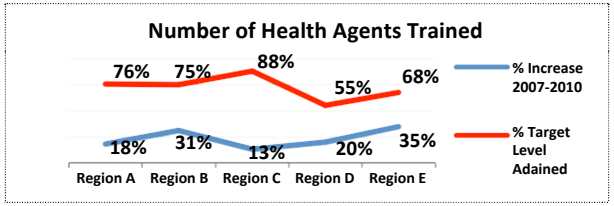
Use of one horizontal and vertical axis is better for the comprehension of a data graphic.

Do not add a second vertical axis

RECOMMENDATIONS FOR GOOD DATA GRAPHICS (5)

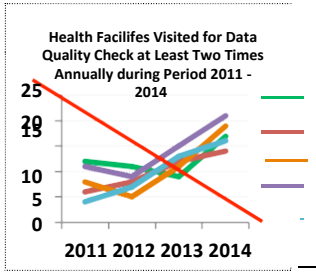
SUMMARY

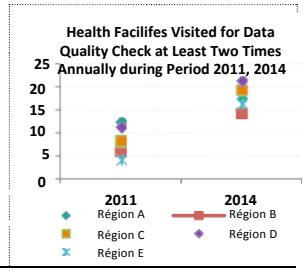
By their very nature, data graphics draw attention. Thus, develop the more meaningful data graphics. However, their excessive use reduces the opportunity to communicate effectively with them.



Data graphics serve to highlight meaningful findings or a very important conclusion

A data graphic should clearly and easily demonstrate the value of the data presented, responding in advance to questions on the statistical meaning and the practical relevance of the findings

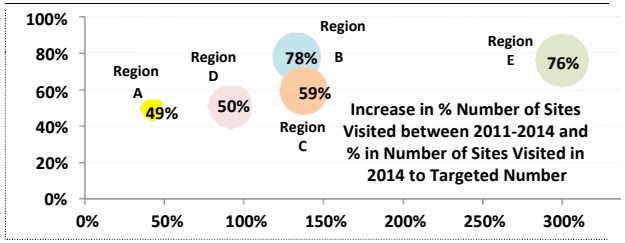




Type of data graphic is appropriate for the presentation of the data

The type of data graphic should be well chosen for the presentation of the data and the comparative analysis undertaken.

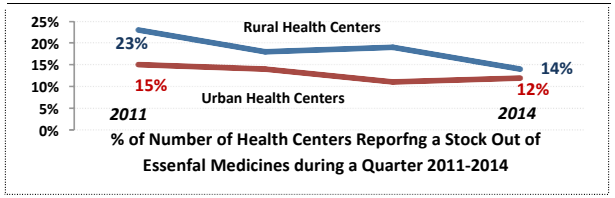
For example, one shows the change of data over the course of time with a line graph, an area graph, a slope graph, or a dot plot



Data graphic has a measure of precision that is practical and appropriate

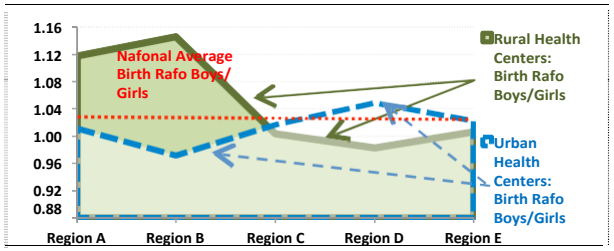
There is not often a need to have label numbers with decimal values. When precision is truly important, choose a type of data graphic that clearly illustrates differences by means of lengths or spread of points along a line.

When precision is less important, use a data graphic that demonstrates the differences by means of areas, such as wedges of a pie chart or areas circles



Context and comparative analysis of data

Comparative analysis of the data – undertaken over the course of time, across programmatic interventions, among targeted subpopulations – facilitates the understanding of significant data findings



Careful considerations of all the graphic elements can lead to a better dissemination of the message to be remembered

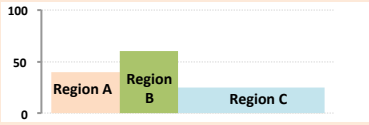
All aspects selected for a data graphic – the type of data graphic, the arrangement of elements, the text, colors, lines – should work together in order to ensure remembrance of the key message

What Would You Like to Show?

✓ COMPARISON among STUDY SUBJECTS

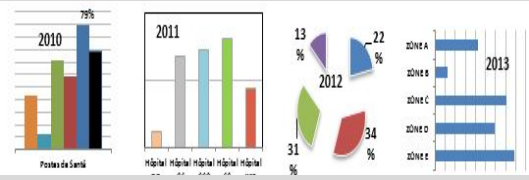
- ***Two Variables for each Study Subject***

➢ **COLUMN CHART of VARIABLE WIDTH**



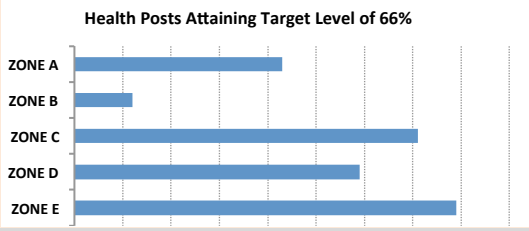
- ***One Variable for each Study Subject***
  - *Large Number of Categories*

➢ **TABLE, TABLE with INTEGRATED GRAPHICAL ELEMENTS**



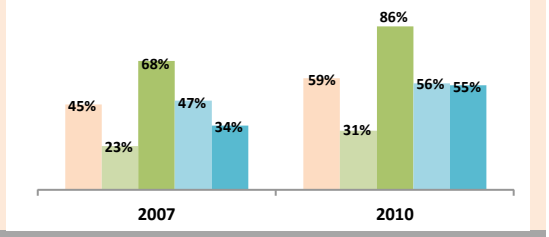
- ***One Variable for each Study Subject***
  - *Small Number of Categories*
    - *Large Number of Study Subjects*

➢ **BAR CHART**



- ***One Variable for each Study Subject***
  - *Small Number of Categories*
    - *Small Number of Study Subjects*

➢ **BAR CHART**

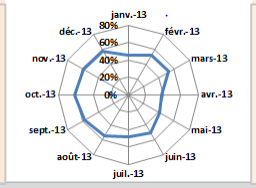


✓ COMPARISON over COURSE of TIME

- ***Large Number of Time Periods***
  - *Data Collected over the Course of a Time Cycle*

➢ **CIRCULAR AREA CHART**

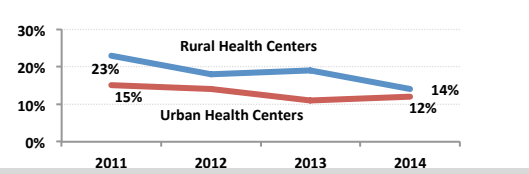
Rate of Timely Submission of Monthly District Reports



January – December 2013

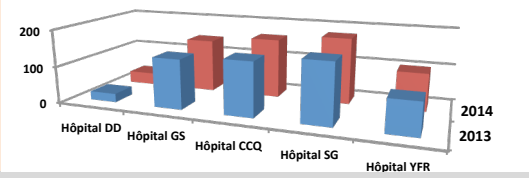
- ***Large Number of Time Periods***
  - *Data NOT Collected over the Course of a Time Cycle*

➢ **LINE CHART**



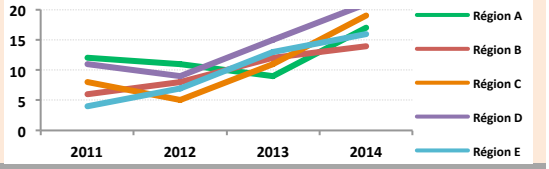
- ***Small Number of Time Periods***
  - *One or a Few Categories*

➢ **COLUMN CHART**



- ***Small Number of Time Periods***
  - *Large Number of Categories*

➢ **LINE CHART**

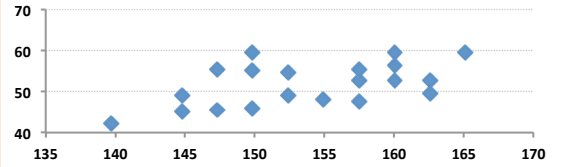


What Would You Like to Show?

✓ STATISTICAL RELATIONSHIP

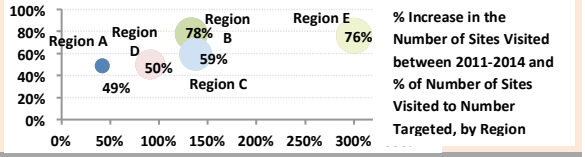
- *Two Variables*

➢ **SCATTER PLOT**



- *Three Variables*

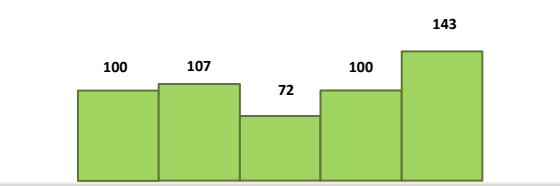
➢ **BUBBLE CHART**



✓ DISTRIBUTION

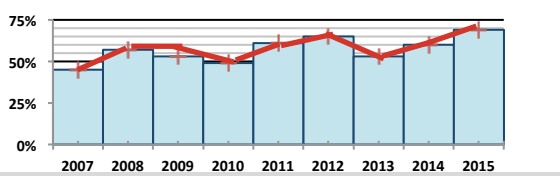
- *One Variable*
  - *Small Number of Data Points*

➢ **HISTOGRAM**



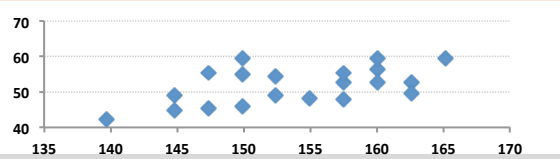
- *One Variable*
  - *Large Number of Data Points*

➢ **LINE HISTOGRAM**



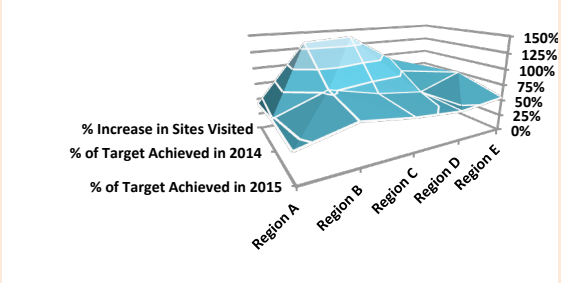
- *Two Variables*

➢ **SCATTER PLOT**



- *Three Variables*

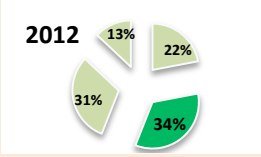
➢ **AREA CHART in THREE DIMENSIONS**



✓ COMPOSITION

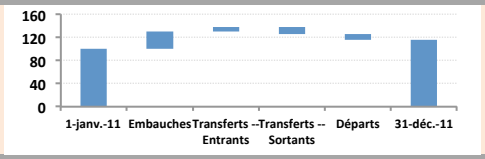
- ***Constant over the Course of Time (Static)***
  - *Proportional Parts of the Whole*

➢ **PIE CHART**



- ***Constant over the Course of Time (Static)***
  - *Additions to the Whole, Subtractions from the Whole*

➢ **WATERFALL CHART**

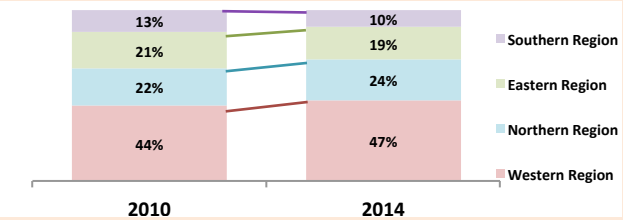


- ***Constant over the Course of Time (Static)***
  - *Parts of Parts*

➢ **100% STACKED COLUMN CHART**

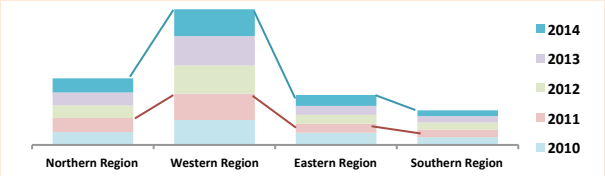
- ***Variable over the Course of Time***
  - *Small Number of Time Periods*
    - *Relative Differences among the Variables are Not Important*

➢ **100% STACKED COLUMN CHART**



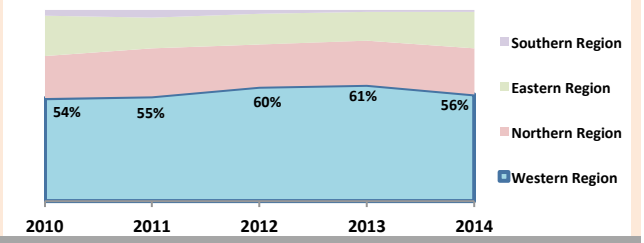
- ***Variable over the Course of Time***
  - *Small Number of Time Periods*
    - *Relative and Absolute Differences among the Variables are Important*

➢ **STACKED COLUMN CHART**



- ***Variable over the Course of Time***
  - *Large Number of Time Periods*
    - *Relative Differences among the Variables are Not Important*

➢ **100% STACKED AREA CHARTS**



- ***Variable over the Course of Time***
  - *Large Number of Time Periods*
    - *Relative and Absolute Differences among the Variables are Important*

➢ **STACKED AREA CHARTS**

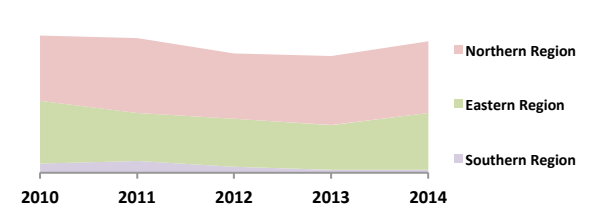
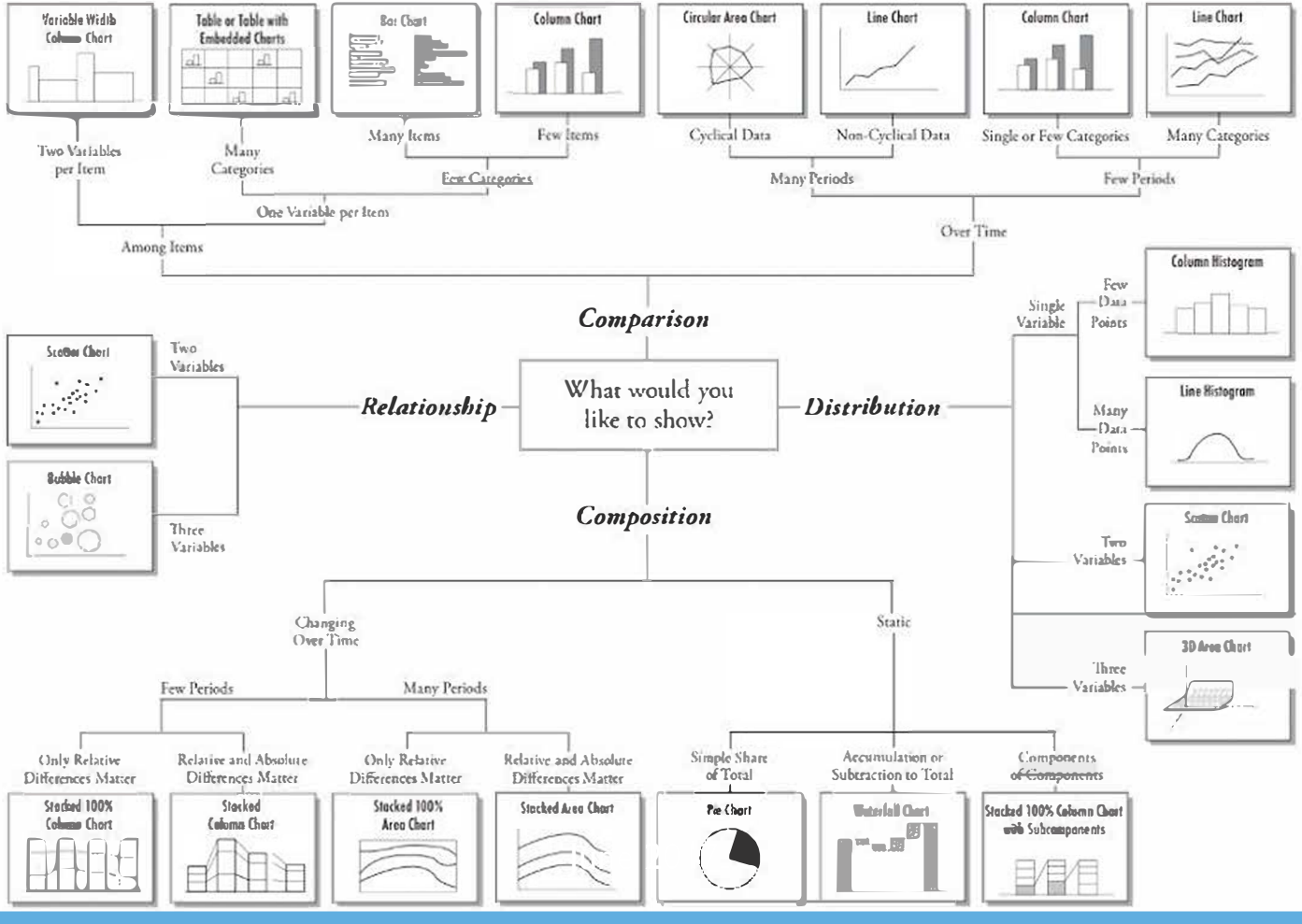


Chart Suggestions—A Thought-Starter

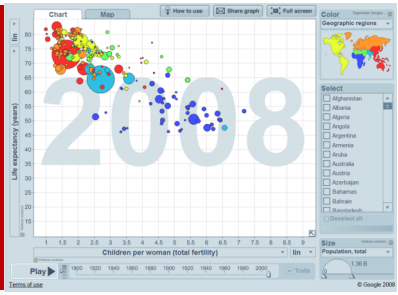


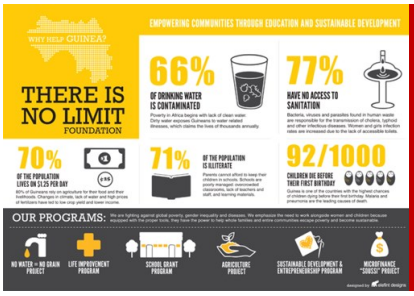
love your data

"They say knowledge is power, **but how do we make knowledge powerful?**"

Resources for learning & loving data visualization

Data visualizations are generated by software, driven by numbers, and can be quickly tweaked and altered.





Infographics are manually drawn; to apply the same look and feel, you have to manually change everything, which can take time.

Identify your audience & context	Find the Story in your Data	Build your Visualization	Disseminate, Share & Use

Creating a Visual Concept

Once you have data or information you want to share using a visualization, there are key questions to think through as you sketch out your concept.

WHO are the stakeholders?
WHAT type of info do they need?
WHERE are they using these data?
WHY do they care about these data?

When answering these questions, think about these key three inputs into your visualization.

Audience: Who is going to look at this visualization? You can't know what data has meaning unless you know who will be interpreting the information.

Data: What information do you have? What story does it tell—correlations, statistics, etc.?

Designer: Who is designing this concept? This could be you, you and your team, or you with an external designer.

Based on the answers to the questions above, determine what kind of visualization would be best: a graph, infographic, video, or another concept.

Your concept will determine the **resources** needed.

Human resources Depending on the visual you're creating, you may be able to manipulate the data yourself or require graphic support from publications staff. Developing visualizations is simplified through various software programs and tools, making data viz a skill anyone with an interest and the time can learn.

Time Consider how much time needs to be budgeted for your visualization(s). For example, are they complex graphics, or simple graphs that could be displayed as a dashboard?

Tools See the reverse side for great software, tools & ideas!

AMANDA MAKULEC / John Snow Inc. amakulec@jsi.com
ERICA NYBRO / Demographic and Health Surveys Program enybro@icfi.com
LIBBY SKOLNIK / JHU-CCP/MLE Project sskolnik@jhuccp.org

DataVizHub.co

Suggested Resources for Learning & Loving Data Visualization

Platforms for creating interesting visualizations*:

DataMarket, http://datamarket.com
Access thousands of data sets from all over the world, including UN and World Bank, or upload your own, and then use the built in tools to create data visualizations.

Tableau, www.tableausoftware.com
Upload your own data and create data visualizations with this very robust software.

Google Fusion Tables, www.google.com/fusiontables
Use Google to easily, quickly make fusion tables with your data.

Dedoose, www.dedoose.com
A mixed-methods analysis software with the capacity to generate compelling visualizations.

Gapminder, www.gapminder.org
Provides a fun interactive platform to explore relationships between variables. Datasets are provided; this platform was made famous by Hans Rosling.

*Note that some open source visualization programs make your data public when you upload. Be sure to read the privacy & data sharing disclaimer.

Platforms for creating infographics:

Piktochart.com, easel.ly and Infogr.am offer templates and design approaches for infographics, if that's the next step for your viz.

Platforms for creating timelines:

Timeline JS (timeline.verite.co) allows you to create free timelines using Excel or Google Docs. Tiki Toki (tiki-toki.com) also creates timelines, but requires you pay for external embed and public views.

Resources to take your data viz to the next level:

Complex Diagrams, http://complexdiagrams.com/
Shows the properties and best uses of visual encodings, i.e. whether it's best to use shape or color to convey your message.

Color Brewer, http://colorbrewer2.org/
Colorbrewer is an online tool designed to help people select good color schemes for maps and other graphics.

Junk Charts http://junkcharts.typepad.com/

Places to find useful health datasets:

Sometimes you might want to create a visualization using secondary data. The following sites are great resources for health-related datasets:

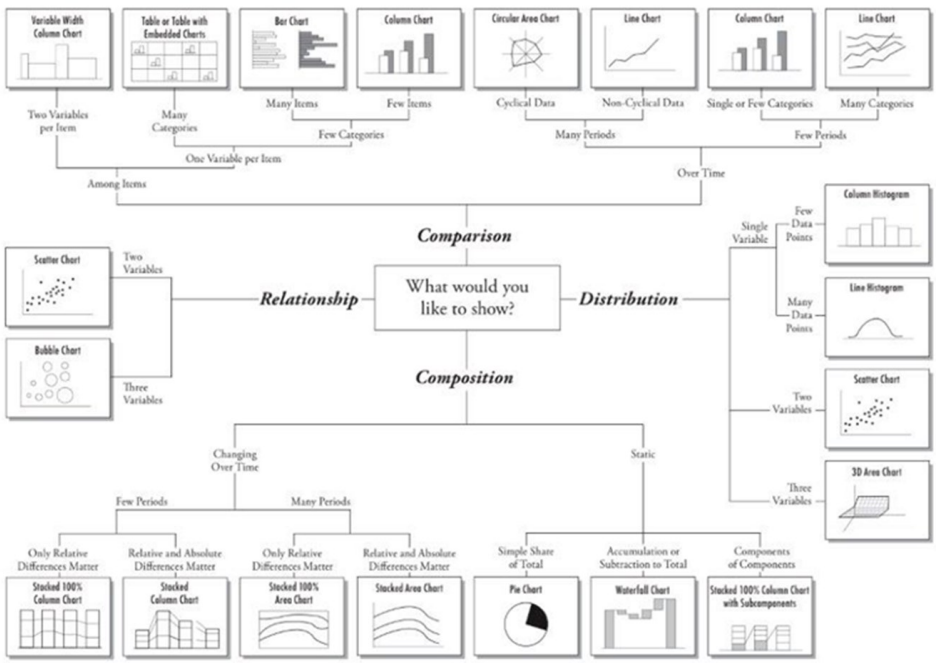
DHS Program—http://www.dhsprogram.com—Contains all DHS datasets .

STATcompiler—http://www.statcompiler.com—Simple interface for downloading select DHS data in table s or visualizations. Also has a mobile app available for Apple, Android & Windows devices for on-the-go stats access!

World Bank Data—http://data.worldbank.org—includes most WHO, DHS, and UN statistics.

Institute for Health Metrics and Evaluation (IHME) Global Health Data Exchange—http://www.healthmetricsandevaluation.org/ghdx—Includes data from all IHME research and others.

AIDSVu.org—Has both state and county level HIV data



Choose-your-own-adventure of charts

Hacking Beautiful Graphs in Excel

Have no fear: if your tool of choice is the ubiquitous MS Excel, you can still create beautiful visualizations.

Ann Emery (http://annkemery.com/excel) has excellent short video tutorials in her for creating interesting graphs using Excel and some creative maneuvering of data, rows, and columns.

Think about how to make your Excel graph or chart sing: **change fonts, color palettes, and use other formatting tricks** to trick people into thinking you used a fancy viz tool.

Think about your choice of chart types: avoid creating pie charts, which can be hard for the human eye to see, and consider horizontal bar charts instead of vertical ones, which give you more space for your axis labels and categories.

Check out the flowchart at left for some inspiration to help you go beyond the bar chart!

Data Visualization Checklists

The next two resources contain checklists of all the criteria you should look for in a good data visualization.

This checklist is meant to be used as a guide for the development of high impact data visualizations. Rate each aspect of the data visualization by circling th most appropriate number, where 2 points means the guideline was fully met, 1 means it was partially met, and 0 means it was not met at all. p/a should not b used frequently, but reserved for when the guideline truly does not apply. For example, a pie chart has no axes lines or tick marks to rate. Refer to the Dat Visualization Anatomy Chart on the last page for guidance on vocabulary.

GUIDELINES

TEXT
Graphs don't contain much text, so existing text must encapsulate your message and pack a punch

6-12 word descriptive title is left---justified in upper left corner

Short titles enable readers to comprehend takeaway messages even while quickly skimming the graph. Rather than a generic phrase, use a descriptive sentence that encapsulates the graph's findings or "so what"? Western cultures start reading in the upper left, so locate title there

Subtle and/or annotations provide additional information

Subtitles and annotations (call-out text within a graph) can add explanatory and interpretive power to a graph. Use them to answer questions a viewer might have or to highlight one of two data points

Text size is hierarchical and readable

Titles are in larger size than subtitles or annotations, which are larger than labels, then by axis labels, and lastly by source information. The smallest text, axis labels, are at least 9 point font size on paper and at least 20 point size on screen

Text is horizontal

Titles, subtitles, annotations, and data labels are horizontal (not vertical or diagonal). Line labels and axis labels can deviate from this rule and still receive full points

Data are labeled directly

Position data labels near the data rather than in a separate legend (e.g., on top of or next to bars or pie slices, and next to lines in line charts).Eliminate/embed legends when possible because eye movement back and forth between the legend and the data can interrupt the brain's attempt to interpret the graph

Labels are used sparingly

Focus attention by removing the redundancy. For example, in line charts, label every other year on an axis

ARRANGE-MENT
Improper arrangement of graph elements can confuse read-ers at best and mislead them at worst. Thoughtful arrangement makes for easier data interpretation

Proportions are accurate

A viewer should be able to take a ruler to measure the length or area of the graph and find that it matches the relationship in the underlying data

Data are intentionally ordered

Data should be displayed in an order that makes logical sense to the viewer. Data may be ordered by frequency counts (e.g., from
greatest to least for nominal categories), by groupings or bins (e.g., histograms), by time period (e.g., line charts), alpha

Axis intervals are equidistant

The spaces between axis intervals should be the same unit, even if the axis interval isn't labeled

Graph is two-dimensional

Avoid three-dimensional displays, bevels, and other distortions

Display is free from decoration

Graph is free from clipart or other illustrations used solely for decoration. Some graphics, like icons, can support interpretation

COLOR

Keep culture-laden color connotations in mind. For example, pink is highly associated with feminine qualities in the US. Use sites like Color Brewer to find color schemes for reprinting in B/W and for color blindness

Color scheme is intentional

Colors should represent brand or intentional choice, not default color schemes. A safe bet for consultants is to use your client's colors. Use online tools to identify brand colors and others that are compatible.

Color is used to highlight key patterns

Action colors should guide the viewer to key parts of the display. Less important or supporting data should be a muted col

Color is legible when printed in black and white

When printed or photocopied in black and white, the viewer should still be able to see patterns in the data

Color is legible for people with colorblindness

Avoid red-green and yellow-blue

Text sufficiently contrasts with background

Combinations when those colors touch one another Black/very dark text against white/transparent background is easiest to read.

LINES

Excessive lines – gridlines, borders, tick marks, axes – can add clutter to a graph, so eliminate them when they are not useful for interpretng data

Grid lines, if present, are muted

Color should be faint gray, not black. Full points if no grid lines are used

Graph does no have a border line

Graph should bleed into the surrounding page or slide rather than being contained by a border

Axes do not have unnecessary tick marks

Tick marks are useful in line graphs (to demarcate each point in time along the y-axis) but unnecessary in bar charts

Graph has one horizontal and one vertical axis

Viewers can best interpret one x-axis and one y-axis, even if one is hidden. Don't add a second y-axis

OVERALL

Graphs will catch a viewer's attention so only visualize the data that needs attention. Too many graphics of unimportant information dilute the power of visualization

Graphs highlights significant finding or conclusion

Graphs should have a "so what?" – either a practical or statistical significance (or both) to warrant their presence

The type of graph is appropriate for data

Data are displayed using a graph type that is appropriate for the relationship within the data. For example, change over time is displayed as a line graph, area chart, slope graph, or dot plot

Graph has appropriate level of precision

Few numeric labels need decimal places. When precision is important, choose a type of graph type that displays differences through length or points along a line (e.g., bar charts, dot plots). When precision is less important, you can use a graph that displays differences through angles or area (e.g., pie charts, circle charts)

Contextualized or comparison data are present

Comparisons – over time, across programs or subgroups of participants, etc. – help the viewer understand the significance of the data

Individual chart elements work together to reinforce the overarching takeaway message

Choices about graph type, text, arrangement, color, and lines should reinforce the take-away message

CHECKLIST for ASSESSING a DATA VISUALIZATION

It is not necessary to check off "YES" for all the questions hereinbelow in order to come up with a very good visualization. Nonetheless, you should take into consideration these aspects concerning content and design of a visualization before you finalize it. If you do not check off a box, make sure that you can justify your reasons for not following these recommendations.

Selection of the Type of Data Graphic

☐ **Does the visualization properly serve to represent the analysis of the data?**

☐ **Would the targeted audience be able to grasp the analysis as represented by the visualisation?**

Simplicity

☐ **Has there been an assessment to determine if features of the visualization are rather unnecessary and could be removed, specifically:**

☐ • Grid lines

☐ • Tick marks

☐ • Horizontal and vertical axes

☐ • Borders

☐ **Are labels presented in as simple a style as possible?**

☐ **Is the data visualisation/data graphic easy to be read and interpreted?**

☐ • Are categories directly and clearly labeled?

☐ • Is the legend properly aligned with the graphic's orientation?

☐ • Are data points labeled properly?

☐ • Is the data graphic clearly titled?

☐ • Are technical names and acronyms of indicators/values defined?

☐ • Are colors and text fonts legible for all viewers?

Clarifying the Message to be Sent on the Data Findings

☐ **Is there need for text to summarize the key message or to articulate the corrective actions to take?**

☐ **Are aspects of pre-attentive attributes strategically used to help clarify the key message?**
(limit to one or two the number of such aspects of pre-attentive attributes)

Honesty

- [ ] **Does the data visualization faithfully represent the data and their findings?**
- [ ] • Do the quantitative values along the Y axis begin at zero?
- [ ] • Do the percentage values of sections comprising a total within a data graphic add up to 100%?
- [ ] • Does the data graphic dispense with using three-dimensional representation?

Additional Questions for Data Visualizations in the form of Infographics, Video, or Dashboard

- [ ] **Does the selected format represent the most appropriate choice?**
- [ ] • Does the format of the data visualization allow for a faithful representation of the data and its transmission of a key message based on that data?
- [ ] • Could there be another more effective format for this data visualization? (for example, there is not always need for an infographic)
- [ ] • Is the data visualization's format appropriate for the expected means to disseminate the information? (examples: Is there a formal plan for disseminating an infographic on the web? Do members of a targeted audience have access to the web for a video posted on YouTube? How could you distribute to the targeted audience information in printed form?

Flow of the Data Message

- [ ] **Is there a title/subtitle/a summary statement that offers a clear key message for a set of data visualizations?**
- [ ] **Do the individual elements (tables, text, photos) within a data visualization complement one another in telling the data message?**
- [ ] **Are the characteristic of elements within a data visualization — font, colors, scale of sizes — consistent in appearance throughout the visualization?**

Overall Design

- [ ] **Is the visualization interesting and pleasing to look at in an aesthetic sense?**
- [ ] **Does the design succeed in reinforcing the effective transmission of the key message?**

Post-Test

The post-test should be given at the end of the final day of the workshop. The purpose is to test the participants on the skills and knowledge gained during the two-day training.

MLE PROJECT: TRAINING in DATA VISUALIZATION

POST-TEST

Please check off your response for each of the following questions
– unless indicated differently by the instructions

1. Data visualization is a way of rebalancing so that information transmission is promoted through visual perception, taking advantage of the powerful capacity of eyes to scan information.

☐ True

☐ False

2. Data visualization is primarily concerned with graphic design.

☐ True

☐ False

3. Charts and graphs developed in the usual way are not considered data visualizations.

☐ True

☐ False

4. The principal point to make (related to findings from a data analysis) should not be directly specified in a data visualization.

☐ True

☐ False

5. From a set of data, there is only one principal point or finding that can be shared with a target audience that can be recognized as a correct and meaningful interpretation.

☐ True

☐ False

POST-TEST (continued)

6. Data visualizations greatly serve to disseminate the sense of the data ("what does it mean?") and to encourage data use for corrective action.

- ☐ True
- ☐ False

7. When presenting a percentage distribution chart, a pie chart is always recommended to be used.

- ☐ True
- ☐ False

8. In a graphical presentation of the data, it is commonly recommended to create a chart in three-dimensional form in order to reinforce its visual impact.

- ☐ True
- ☐ False

9. In a graphical presentation created in several colors, one should consider the potential problem posed by their different combinations for those who are colorblind.

- ☐ True
- ☐ False

10. In a graphical presentation created in several colors, one does not need to consider the cultural associations of colors.

- ☐ True
- ☐ False

POST-TEST (continued)

11. Using the numbers «1» to «4», put in chronological order the following four steps concerning the design process for a data visualization.

- [ ] Analyze the data to determine if there can be found a pattern, a trend, a comparison, or outlier values; from these observations, seek out a meaningful interpretation of the data
- [ ] Sketch out with pencil and paper several ideas that come to you to assess the best approach for transmitting the data analysis findings with a data visualization, recognizing that visual approaches often make information more understandable and compelling than explanatory text
- [ ] Determine your target audiences and stakeholders with whom the data visualization will be shared and identify what each of them should of them should understand from the data and data finding
- [ ] Before the dissemination of a data visualization, customize it for each target and for each method of presentation (report charts are done differently than images for an infographic or a video), recognizing that the necessary data needs to be updated

12. Indicate with a check mark if each of the following items is **True** regarding who are the different target audiences for a data visualization about program data.
[There may be more than one correct response among the four items.]

- [ ] Program managers
- [ ] Program service providers
- [ ] Journalists
- [ ] Researchers

13. Indicate with a check mark if each of the following items is **True** regarding the considerations to keep in mind for each of the different target audiences when developing a data visualization about program data.
[There may be more than one correct response among the four items.]

- [ ] Average level of skills in numeracy (math calculations)
- [ ] Average level of skills in program design
- [ ] Average level of skills in graphic design
- [ ] Average level of skills in data interpretation (data analysis)

POST-TEST (continued)

14. Indicate with a check mark if each of the following items is **True** regarding a data visualization for a program.
[There may be more than one correct response among the four items.]

- [ ] It is a presentation of complex ideas in the form of a story
- [ ] It is a presentation of findings with proper consideration for graphic display
- [ ] It is a presentation of information done in a way to facilitate comprehension
- [ ] It is a presentation of data done to engage with a certain target audience and to instill a response for corrective action

15. Indicate with a check mark if each of the following items is **True** regarding the means for enhancing transmission of meaningful findings by a data visualization.
[There may be more than one correct response among the four items.]

- [ ] Colors are carefully selected
- [ ] All text font is boldfaced
- [ ] Chart legend is placed to the side in order to free up space in the middle
- [ ] Chart title summarizes data findings and is written to highlight their significance

16. Indicate with a check mark if each of the following items is **True** regarding the development of a data visualization.
[There may be more than one correct response among the four items.]

- [ ] Chart axis showing values of magnitude should begin at zero
- [ ] Tick marks should be removed from axes, unless they are necessary
- [ ] Grid lines should be clearly visible in a chart
- [ ] Label and legends should have borders to better encapsulate their words

POST-TEST (continued)

17. Indicate with a check mark if each of the following items is **True** regarding the development of an infographic. *[There may be more than one correct response among the four items.]*

- [ ] Series of images used in an infographic should run from left to right
- [ ] Expected understanding of an infographic by the target audience should be considered
- [ ] Infographic should be developed by a team of analysts and graphic designers
- [ ] Interpretation to be transmitted by an infographic should be first formulated by means of storyboarding

18. Match each term with its definition found to the right.
[Document your choices by writing the alphabet code in the box in front of each term.]

[ ]	Pattern	**A**	Detection of similarities and differences among data points in a data set
[ ]	Trend	**B**	Detection of the minimum and maximum values in a data set
[ ]	Comparison	**C**	Detection of a regular and repeated relationship in the distribution of data points in a data set
[ ]	Range	**D**	Detection of a distribution of data points that conforms to a certain arrangement over the course of time

19. Match each chart type with the reference found to the right concerning the type of interpretation of data to be presented. *[Document your choices by writing the alphabet code in the box in front of each chart type.]*

[ ]	Scatter Plot	**A**	Comparison of total fertility rates among five countries
[ ]	Line Graph	**B**	Distribution by age in a country over the course of time
[ ]	Pie Chart	**C**	Distribution by sex where 60% of the national population is female
[ ]	Bar Chart	**D**	Total fertility rate with respect to prevalent utilization rate of contraceptives in each of twenty countries
[ ]	100% Area Chart	**E**	Fertility rate trend over the course of time in each of several countries

POST-TEST (continued)

20. Match each type of technical specialization with the reference found to the right of tasks involved in the development of a data visualization.
[Document your choices by writing the alphabet code in the box in front of each type of specialization.]

☐	Data Quality Monitoring	**A**	Presentation of an interpretation of the data in a way that is appropriate for the target audience
☐	Technical Programming	**B**	Collection of accurate and complete data for the visualization and quality control of the data interpretation in the visualization itself
☐	Communications	**C**	Quality control on the reporting of programming carried out in relation to the visualization's interpretation of the data that is being transmitted
☐	Computer Programming	**D**	Identification and refinement of messages to be remembered by target audiences from the data visualization
☐	Graphic Arts	**E**	Adaptation of a data visualization so that it can be transmitted on another platform (for example, on the Web)

21. Match each type of data visualization with its definition found to the right.
[Document your choices by writing the alphabet code in the box in front of each term.]

☐	Infographics	**A**	Means for presenting dynamically and specifically quantitative data on demand with customized software
☐	Charts and Graphs	**B**	Means for broadly transmitting an interpretation of quantitative and qualitative data through a series of static images
☐	Videos, Multimedia	**C**	Means for presenting easily and precisely an interpretation of the data in geometric and static forms
☐	Interactive Web Sites	**D**	Means for dynamically transmitting an interpretation of the data through animated images and the use of narrative and sound

Additional Resources

The following pages direct participants to websites, resources, and tools that will further guide them with creating appealing and meaningful data visualizations.

GLOBAL HEALTH LEARNING CENTER

Data Visualization - An Introduction

https://www.globalhealthlearning.org/course/data-visualization-brave-new-world

BETTER EVALUATION

http://betterevaluation.org/plan/describe/visualise_data

Dr. Edward Tufte

"The Visual Display of Quantitative Information"; "Visual Explanations"; "Beautiful Evidence"

https://www.youtube.com/watch?v=Th_1azZA2OY "Beautiful Evidence"

Hans Rosling (YouTube, TED Talks)

"200 Countries, 200 years , 4 Minutes": https://www.youtube.com/watch?v=jbkSRLYSojo

"Joy of Stats": https://vimeo.com/18477762

Andrew Kirk (journalist)

"8 Hats of Data Visualization Design"

https://vimeo.com/44886980

Alberto Cairo (journalist; MOOC on data visualization)

"The Functional Art"

http://www.thefunctionalart.com/

Cole Nussbaumer

Storytelling with Data

Stephanie Evergreen

Consulting Site: http://stephanieevergreen.com/

"Presenting Data Effectively: Communicating Your Findings for Maximum Impact" (Sage Publisher, October 2013)

Ann K. Emery

Blog: http://annkemery.com/

http://annkemery.com/essentials/

Jon Schwabish (Urban Institute, Congressional Budget Office)

Blog: http://blog.visual.ly/author/jon-schwabish/

https://policyviz.com/podcast/episode-58-the-financial-times-graphic-desk/

David Giard

Effective Data Visualization: https://www.youtube.com/watch?v=nP6qWhOkha4

Nancy Duarte
Diagrammer (4,000 customizable diagrams to download for free with which to enhance your PowerPoint presentations): http://www.duarte.com/diagrammer/

8 Great Books about Data Visualisation (Andy Cotgreave)"
http://www.tableau.com/about/blog/2013/7/list-books-about-data-visualisation-24182

Steven Few (business analytics, dashboards)
"Now You See It: Simple Visualization Techniques for Quantitative Analysis" (BOOK)

CONNECT

www.DataVizHub.co

Appendix I: Powerpoint Presentations

This appendix contains two sets of Powerpoint presentations to be used by the workshop facilitator. Each set of presentations represents one day of training for a total of two training days. The presentations should be used as a facilitation guide to walk participants through the training material.

There are interactive activities dispersed throughout the presentations to get engagement and input from participants and to ensure they understand the material. In addition, there are breaks placed throughout the day for lunch and coffee/tea.

Training in Data Visualization

Day 1, Morning –

MLE

Overview of the Training (1)

Days 1 & 2:

- Presentation of technical content
- Interactive activities for the participants, working in groups or in pairs
- Recap of topics covered at each day's end
- Participant evaluations of the morning and afternoon sessions
- Pre-test at the start of Day 1 and Post-test at Day 2's conclusion

MLE

Objectives of this Training

At the end of this training, participants will be able to:

- Understand the concepts, basic principles and good practices in data visualization regarding the development of effective communication materials (reports, infographics, conference posters, dashboards, web-based videos, film)
- Gain a heightened sensitivity to the context and characteristics of key audiences targeted to receive and respond to data findings from health surveys

MLE

At the End of this Session, Participants will be Able to:

1. Define the term: “data visualization”
2. Describe the organization of this training and expected outcomes
3. Categorize their earlier experiences in data visualization
4. Enhance these experiences through assigned taks throughout the course of this training

MLE

Building a Base:
Definition of Data Visualization

Each of the Following is Correct:

- ✓ Process for presenting information that draws inspiration from the ***science of human perception*** and is sensitive to ***cultural and organizational contexts*** regarding need for and interpretation of information
- ✓ Approach in transmitting a ***key message*** based on complex data findings by means of a ***visual presentation*** designed to be ***accessible and persuasive***
- ✓ Strategy for creating communications that ***foster a compelling response*** to data findings ***among target audiences***

MLE

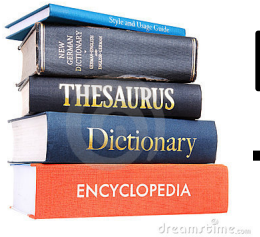
Building a Base:
Need to Give Sense to Data

" Being a member of the global community, each of us who communicates on the Web contributes our daily share of 36,000,000,000,000 (36 trillion) words …. "

- Clive Thompson, Wired

Ashatu Abdullahi, who enters data on family planning services, looks for a record among a mountain of questionnaires kept in an office of the MLE Project, located in Abuja, the national capital of Nigeria.

MLE



Building a Base:
Terms Related to Data Visualization (1)

Pattern

- Analysis of data points to determine if there is a regular and repeated relationship in their distribution

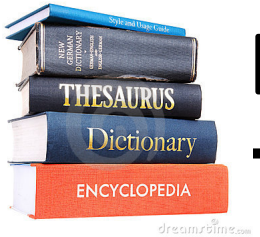
Trend

- Analysis of data points to determine if there is a distribution that follows a certain pattern over the course of time

Comparison

- Analysis of data points to determine if there are similarities and differences

MLE



Building a Base:
Terms Related to Data Visualization (2)

Range

- Analysis of data points to determine the difference between the minimum and maximum values

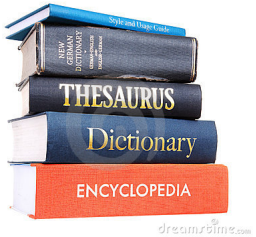
Outlier

- Data point that lies far away from the other data points or does not conform to the same pattern as the others; extreme case

Surprise:

- Unexpected finding based on analysis of the data; conclusion that is unforeseen, as based on data sources

MLE



Building a Base:
Four Principles of Data Visualization

1. Selection of an Appropriate Graphical Image
2. Simplicity in Presentation Style
3. Selective Choice of Pre-Attentive Attributes
4. Honesty

MLE

Process: Developing a Data Visualization

1. Identify your audience & context
- Who are the stakeholders?
- What do they need to know?

2. Find the story in your data
- What information do you have?
- Are there identifiable patterns, trends, surprises, relationships, successes or failures?

3. Build your Visualization
- Sketch first!
- Choose the right type of chart for your data
- Simplify
- Use pre-attentive attributes to highlight your story

4. Disseminate, share, and use!
- How will this visualization be used?
- What is the ideal format?
- Good data visualization ensures that data are read, understood, and used for evidence-based decision-making.

MLE

What Data are You Sharing?

Table 4.3. Current use of contraception

Percent distribution of women by type of contraceptive method currently used by wealth quintile and city. Kenya 2010, 2012.

	Baseline family planning use				Mid-term family planning use			
	Modern[1]	Traditional[2]	Not using	Total	Modern[1]	Traditional[2]	Not using	Total
Nairobi								
Poorest	35.9	3.8	60.4	100.0	42.7	7.9	49.4	100.0
Poor	43.9	3.3	52.8	100.0	56.8	7.6	35.7	100.0
Middle	52.5	4.3	43.3	100.0	55.8	7.3	36.9	100.0
Rich	43.6	4.4	52.0	100.0	47.5	7.2	45.2	100.0
Richest	40.5	4.5	55.1	100.0	38.9	7.9	53.1	100.0
Overall	43.7	4.1	52.3	100.0	48.3	7.6	44.1	100.0
Overall - in union	58.0	4.9	37.2	100.0	62.0	8.3	29.7	100.0
Mombasa								
Poorest	24.1	4.7	71.2	100.0	42.1	8.2	49.7	100.0
Poor	31.0	3.2	65.8	100.0	34.0	9.7	56.3	100.0
Middle	36.1	5.5	58.4	100.0	33.1	8.7	58.2	100.0
Rich	30.6	3.1	66.3	100.0	31.7	10.7	57.6	100.0
Richest	28.0	5.0	67.0	100.0	29.5	9.4	61.0	100.0
Overall	29.4	4.4	66.3	100.0	34.2	9.3	56.5	100.0
Overall - in union	41.7	6.7	51.7	100.0	44.2	11.4	44.4	100.0
Kisumu								
Poorest	44.3	3.0	52.7	100.0	55.5	2.7	41.8	100.0
Poor	44.9	3.5	51.6	100.0	61.0	1.6	37.4	100.0
Middle	41.8	4.4	53.8	100.0	63.9	5.3	30.8	100.0
Rich	47.6	2.4	50.1	100.0	43.5	7.3	49.1	100.0
Richest	42.7	5.0	52.3	100.0	54.3	8.6	37.1	100.0
Overall	44.4	3.6	52.1	100.0	55.5	5.1	39.4	100.0
Overall - in union	52.3	4.6	43.2	100.0	62.8	5.8	31.5	100.0

1 Modern methods include male/female sterilization, daily pill, IUD, injectables, male condom, female condom, LAM/breastfeeding and E-pill

2 Traditional methods include periodic abstinence, withdrawal, and standard days/safe days/cycle beads

Source: Mid-Term Study in Kenya. © 2013 MLE Project

MLE

Why Data Visualization?

Make Data Accessible for Decision Making

Source: Chris Lysy, freshspectrum.com

MLE

Interactive Activity #1 (10 minutes)

- Each participant is asked to consider her/his own skill level in data visualization and to select one of four following classifications. Are you:
  - walking on your hands? 
  - striding? 
  - running? 
  - flying like a bird 
- As a group, create a chart with all of your individual selections

MLE

Interactive Activity #2 (25 minutes)

- Choose a partner/partners to form a pair/group in order to undertake a number of interactive activities over the course of the training
- Create a method for evaluating the quality of instruction for morning and afternoon sessions throughout the training
- Each pair/group is asked to use that method just before the lunch break and just before the end of the day
- In addition, each pair/group is asked to create a chart to collectively present these evaluations

MLE

Process: Developing a Data Visualization

1. Identify your audience & context
- Who are the stakeholders?
- What do they need to know?

2. Find the story in your data
- What information do you have?
- Are there identifiable patterns, trends, surprises, relationships, successes or failures?

3. Build your Visualization
- Sketch first!
- Choose the right type of chart for your data
- Simplify
- Use pre-attentive attributes to highlight your story

4. Disseminate, share, and use!
- How will this visualization be used?
- What is the ideal format?
- Good data visualization ensures that data are read, understood, and used for evidence-based decision-making.

MLE

Process:
Step 1

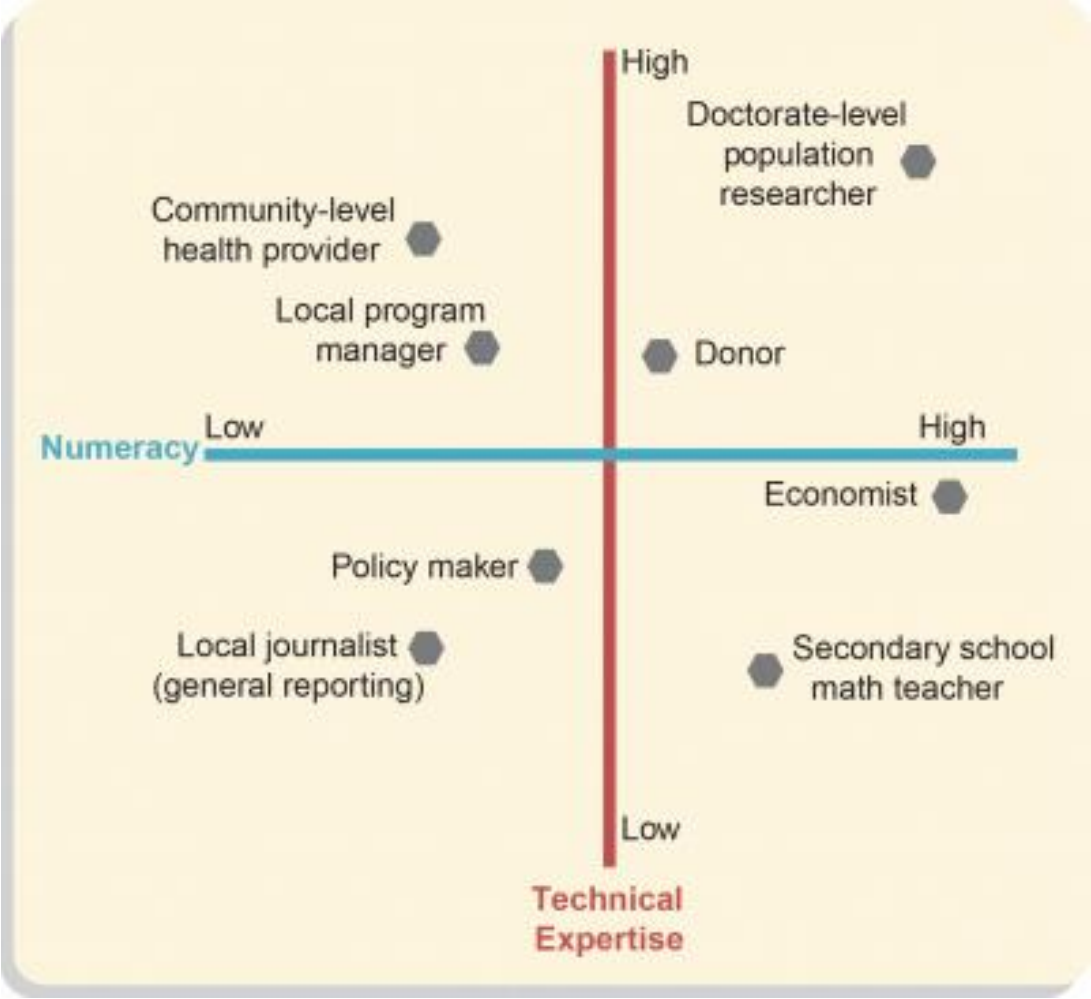
Identification of Target Audiences and Understanding the Context of their Informational Needs

At the end of this session, participants will be able to:

- List the important characteristics to consider before developing a data visualization
- Respond to questions on case studies, showing a good understanding of the context for communications with target audiences

MLE

Matrix of Target Audiences for Information on Health Programs



MLE

Target Audiences: Potential Use of Different Types of Information (1)

- Health Program Policy Makers
  - Overall levels of program inputs and outputs
  - Level of success in meeting performance targets
  - Overall outcomes from health services delivery
  - Levels of efficiency in health service delivery and in program administration
- Health Program Managers
  - Performance measures at different levels (health service delivery, program administration, geographic zones) related to inputs, outputs, adherence to procedures, goal achievements
- Target Populations for Health Services
  - Description, means of access, and benefits of health services offered locally

MLE

Target Audiences: Potential Use of Different Types of Information (2)

- Media
  - Overall levels of program inputs and outputs
  - Level of success in meeting performance targets
  - Outcomes and benefits from health services delivery
  - Levels of efficiency in health service delivery and in program administration
- Donors
  - Overall levels of program inputs and outputs
  - Levels of success in meeting performance targets and in adhering to procedures
  - Outcomes and benefits from health services delivery
  - Levels of efficiency in health service delivery and in program administration
  - Performance level differences by geographic zone

MLE

Coffee / Tea Break (15 minutes)

Interactive Activity #3 (20 minutes)

- Separate yourselves into pairs/groups in order to determine:
  - Types of target audiences that they themselves have attempted to reach in their line of work
  - Levels of skills in numeracy and data analysis of those in each of type of target audience
  - Essential considerations in designing a data visualization for a good response from each of these target audience types

MLE

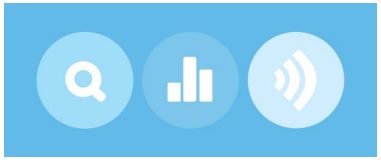
Target Audiences: Potential Use of Different Types of Information (1)

- Health Program Policy Makers
  - Overall levels of program inputs and outputs
  - Level of success in meeting performance targets
  - Overall outcomes from health services delivery
  - Levels of efficiency in health service delivery and in program administration

MLE

Target Audiences: Potential Use of Different Types of Information (2)

- Health Program Policy Makers
  - Overall levels of program inputs and outputs
  - Level of success in meeting performance targets
  - Overall outcomes from health services delivery
  - Levels of efficiency in health service delivery and in program administration



- ✓ Monitoring of performance in order to assess the continuation of financial support and at what level for health programming
- ✓ Improved knowledge on what works better in health programming
- ✓ Advocacy with other policy makers, using program outcomes

MLE

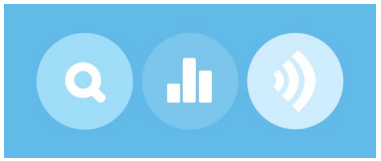
Target Audiences: Potential Use of Different Types of Information (3)

- Health Program Managers
  - Performance measures at different levels (health service delivery, program administration, geographic zones) related to inputs, outputs, adherence to procedures, goal achievements

MLE

Target Audiences: Potential Use of Different Types of Information (4)

- Health Program Managers
  - Performance measures at different levels (health service delivery, program administration, geographic zones) related to inputs, outputs, adherence to procedures, goal achievements

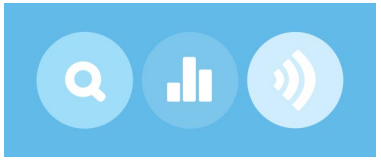


- ✓ Performance monitoring for identifying and correcting shortcomings and recognizing and encouraging successes
- ✓ Management of resources and their redeployment, as needed
- ✓ Analysis of operational and contextual reasons for the levels of delivery and use of health services
- ✓ Advocacy with health program policy makers, using program outcomes

MLE

Target Audiences: Potential Use of Different Types of Information (5)

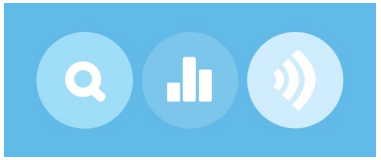
- Target Populations for Health Services
  - Description, means of access, and benefits of health services offered locally



MLE

Target Audiences: Potential Use of Different Types of Information (6)

- Target Populations for Health Services
  - Description, means of access, and benefits of health services offered locally

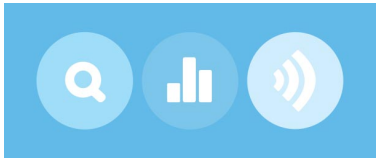


- ✓ Knowledge of health services that are available and accessible
- ✓ Recognition of the value in using health services
- ✓ Improved understanding on how to better safeguard good health for oneself, one's family, and one's community

MLE

Target Audiences: Potential Use of Different Types of Information (7)

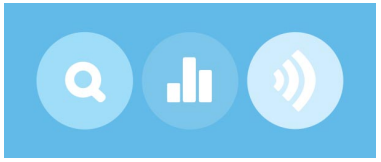
- Media
  - Overall levels of program inputs and outputs
  - Level of success in meeting performance targets
  - Outcomes and benefits from health services delivery
  - Levels of efficiency in health service delivery and in program administration



MLE

Target Audiences: Potential Use of Different Types of Information (8)

- Media
  - Overall levels of program inputs and outputs
  - Level of success in meeting performance targets
  - Outcomes and benefits from health services delivery
  - Levels of efficiency in health service delivery and in program administration



- ✓ Investigative journalism to broadcast to the general public what works and does not work in health services delivery
- ✓ Promotion of good practices to better safeguard health
- ✓ Dissemination of information that is of key interest to the media's audience

MLE

Target Audiences: Potential Use of Different Types of Information (9)

- Donors
  - Overall levels of program inputs and outputs
  - Success in meeting performance targets and adhering to procedures
  - Outcomes and benefits from health services delivery
  - Efficiency in health service delivery and in program administration
  - Performance level differences by geographic zone

MLE

Target Audiences: Potential Use of Different Types of Information (10)

- Donors
  - Overall levels of program inputs and outputs
  - Success in meeting performance targets and adhering to procedures
  - Outcomes and benefits from health services delivery
  - Efficiency in health service delivery and in program administration
  - Performance level differences by geographic zone
    - ✓ Monitoring of performance in order to assess the continuation of financial support and at what level for health programming
    - ✓ Improved knowledge on what would work better in health programming
    - ✓ Dissemination of information that is of key interest to donors

MLE

Objectives of This Session

At the end of this session, participants will be able to review data tables in order to identify a finding that would be clear and relevant to a target audience:

1. Table from an imaginary case study
2. Tables from the field studies of the MLE Project in Kenya

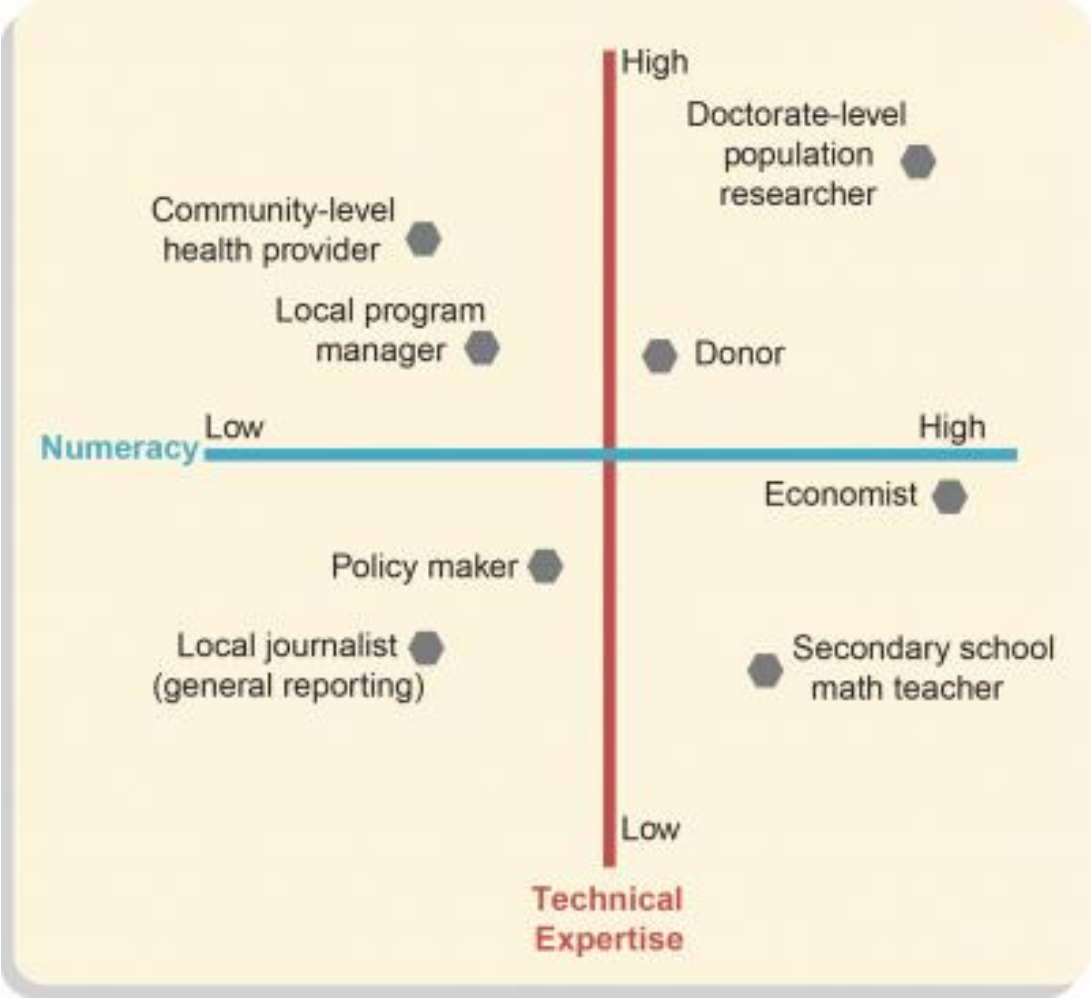
MLE

Interactive Activity #4 (15 minutes)

- Separate yourselves into pairs/groups in order to analyze a case study
  - Identify who are the target audiences
  - Determine the characteristics of these target audiences concerning the context in which they would respond well to data results
  - Specify the type of information and visualizations to be delivered to each of these target audiences
  - Provide reasons for your given responses

MLE

Matrix of Target Audiences for Information on Health Programs



MLE

Interactive Activity #4 (continued)

(15 minutes)

	Frequency of Attendance at Births in the Eastern Region during 2014
Doctors	**42.0 %**
Traditional Midwives	**41.0 %**
Trained Midwife / Nurse / Female Health Agent	**10.2 %**
Family Member	**6.2 %**
No One	**.2 %**
Lack of Information	**.4 %**

MLE

Interactive Activity #4 (continued)

(15 minutes)

Frequency of Attendance at Births in the Eastern Region during 2014

Doctors	42.0 %	52.2 %
Trained Midwife / Nurse / Female Health Agent	10.2 %	
Traditional Midwives	41.0 %	47.4 %
Family Member	6.2 %	
No One	.2 %	
Lack of Information	.4 %	

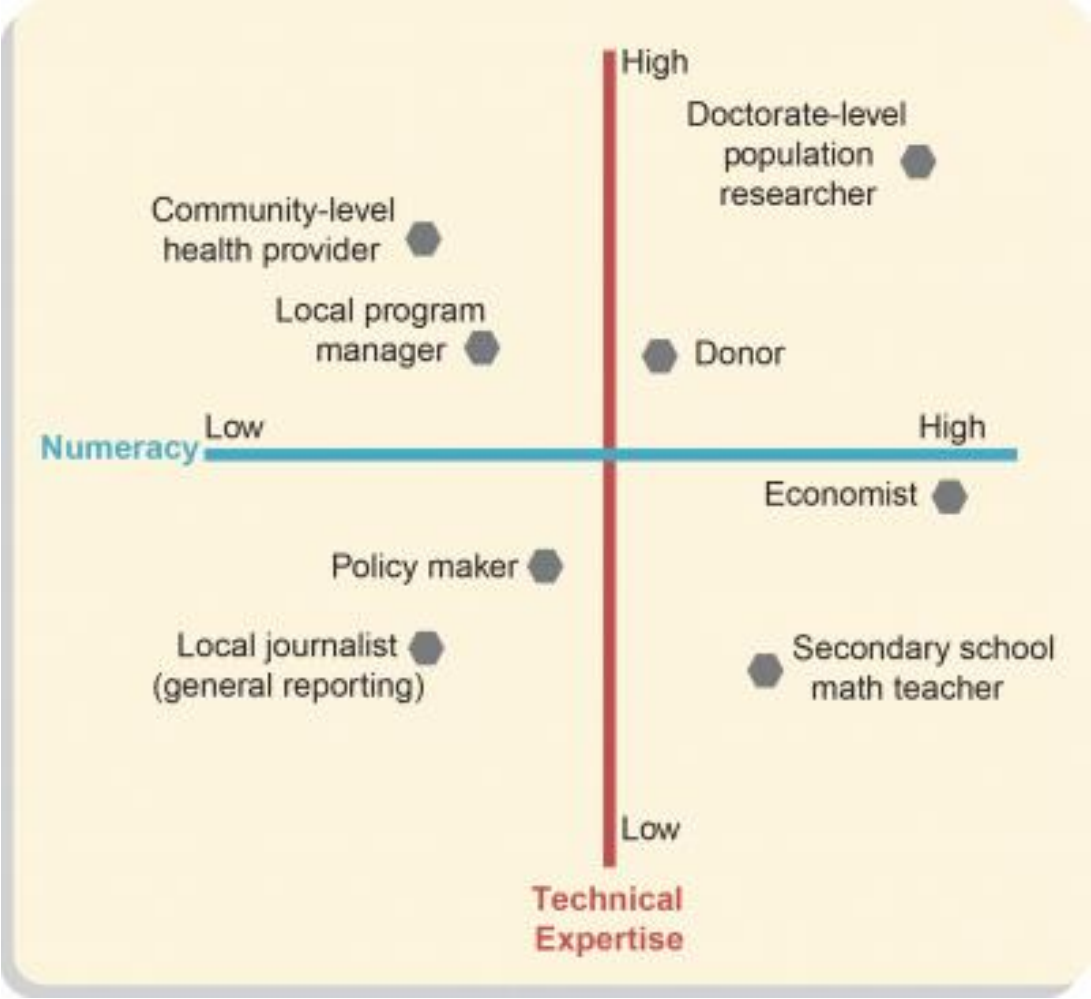
MLE

Interactive Activity #5 (45 minutes)

- Remaining in pairs/groups, analyze a data table from the MLE Project
  - Identify who are the target audiences
  - Determine the characteristics of these target audiences concerning the context in which they would respond well to data results
  - Specify the type of information and visualizations to be delivered to each of these target audiences
  - Provide reasons for your given responses

MLE

Matrix of Target Audiences for Information on Health Programs



MLE

Distribution (in percentage) of women at KENYA MLE Project sites in a relationship, according to their current usage of a contraceptive method at the time of the baseline survey (2010) and end-line survey (2014)

% SURVEYED WOMEN in a RELATIONSHIP Baseline Survey Results End-Line Survey Results	USING ANY METHOD	Any Modern Method	Any Traditional Method	*Non-User*
NAIROBI	47.8 61.6	43.6 54.8	4.2 6.8	*52.2* *38.4*
MOMBASA	33.8 47.1	29.3 43.8	4.5 3.3	*66.3* *52.9*
KISUMU	47.9 62.5	44.4 58.7	3.5 3.8	*52.1* *37.5*
MACHAKOS	53.3 64.6	45.3 57.8	8.0 6.8	*46.8* *35.4*
KAKAMEGA	48.5 55.4	46.1 53.8	2.4 1.6	*51.5* *44.5*

Examples of Data from the Field Studies Carried out by the MLE Project

Distribution (in percentages) of Women at MLE Kenya Project Sites According to their Knowledge (Spontaneous Knowledge or After Prompting) of Contraceptive Methods at the Time of MLE Project Baseline Survey (2010) & End-line Survey (2014)

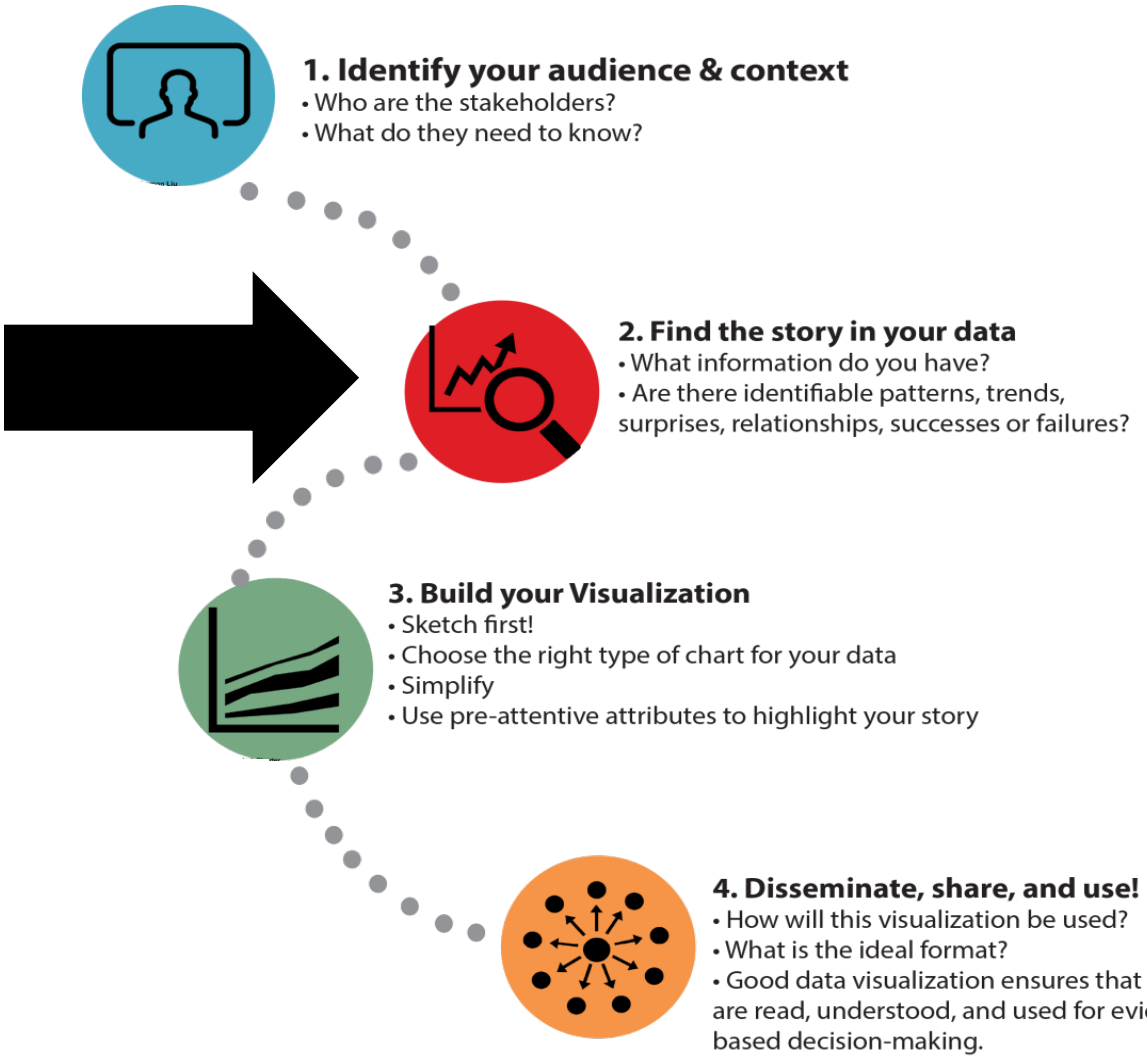
	NAIROBI		MOMBASA		KISUMU	
	Project Baseline 2010	Project End-Line 2014	Project Baseline 2010	Project End-Line 2014	Project Baseline 2010	Project End-Line 2014
	%	%	%	%	%	%
ANY MODERN METHOD	**98.1**	**100.0**	**98.3**	**99.8**	**99.0**	**99.9**
Female Sterilization	63.4	90.8	71.1	80.0	74.8	86.0
Male Sterilization	50.0	76.5	56.7	62.8	52.7	65.6
IUD	82.8	97.1	78.0	95.4	83.3	96.4
Injectables	95.4	99.3	93.4	98.3	96.5	99.4
Implants	81.4	98.0	78.9	94.7	87.0	98.1
Daily Pills	94.9	99.4	94.4	96.8	95.0	99.4
Emergency Contraception	57.7	79.4	44.5	57.6	57.6	67.3
Male Condom	96.7	99.8	97.5	98.7	98.4	99.8
Female Condom	87.1	96.9	74.9	85.7	87.8	95.2
LAM Breastfeeding	49.7	76.7	40.2	77.3	46.8	52.6
Std Days Method, CycleBeads	79.2	65.5	78.9	46.3	78.0	45.7
***TRADITIONAL** -- Withdrawal*	*56.9*	*87.2*	*62.6*	*69.4*	*56.2*	*69.7*

Examples of Data from the Field Studies Carried out by the MLE Project

Distribution (in percentages) of Women at MLE Kenya Project Sites According to their Knowledge (Spontaneous Knowledge or After Prompting) of Contraceptive Methods at the Time of MLE Project Baseline Survey (2010) & End-line Survey (2014)

	MACHAKOS		KAKAMEGA	
	Project Baseline 2010	Project End-Line 2014	Project Baseline 2010	Project End-Line 2014
	%	%	%	%
ANY MODERN METHOD	**98.9**	**100.0**	**98.4**	**100.0**
Female Sterilization	83.8	89.4	76.6	93.6
Male Sterilization	58.3	57.2	65.7	74.0
IUD	85.0	93.8	87.3	97.8
Injectables	95.8	98.3	95.9	99.8
Implants	86.3	96.3	84.0	98.6
Daily Pills	95.5	98.0	94.2	99.8
Emergency Contraception	52.8	51.2	51.9	63.7
Male Condom	98.0	99.7	97.4	99.9
Female Condom	80.0	86.6	90.5	95.2
LAM Breastfeeding	37.7	41.2	54.5	63.2
Std Days Method, CycleBeads	87.5	37.0	81.6	34.6
***TRADITIONAL** -- Withdrawal*	*62.6*	*67.4*	*63.1*	*73.5*

Process: Developing a Data Visualization



MLE

Process:
Step 2

Program for This Session

Learning Activity 1 – Brief overview of important terms; as an exercise, participants match one of these terms to a conclusion to be drawn from each one of a set of data charts (30 minutes)

Learning Activity 2 – Brief overview of the process for carrying out storyboarding in order to formulate a key message based on an analysis of the data; as an exercise, participants will undertake this process, using a set of data charts provided to them (60 minutes)

MLE

Identification of a Key Message to be Transmitted that has been Formulated from an Analysis of the Data

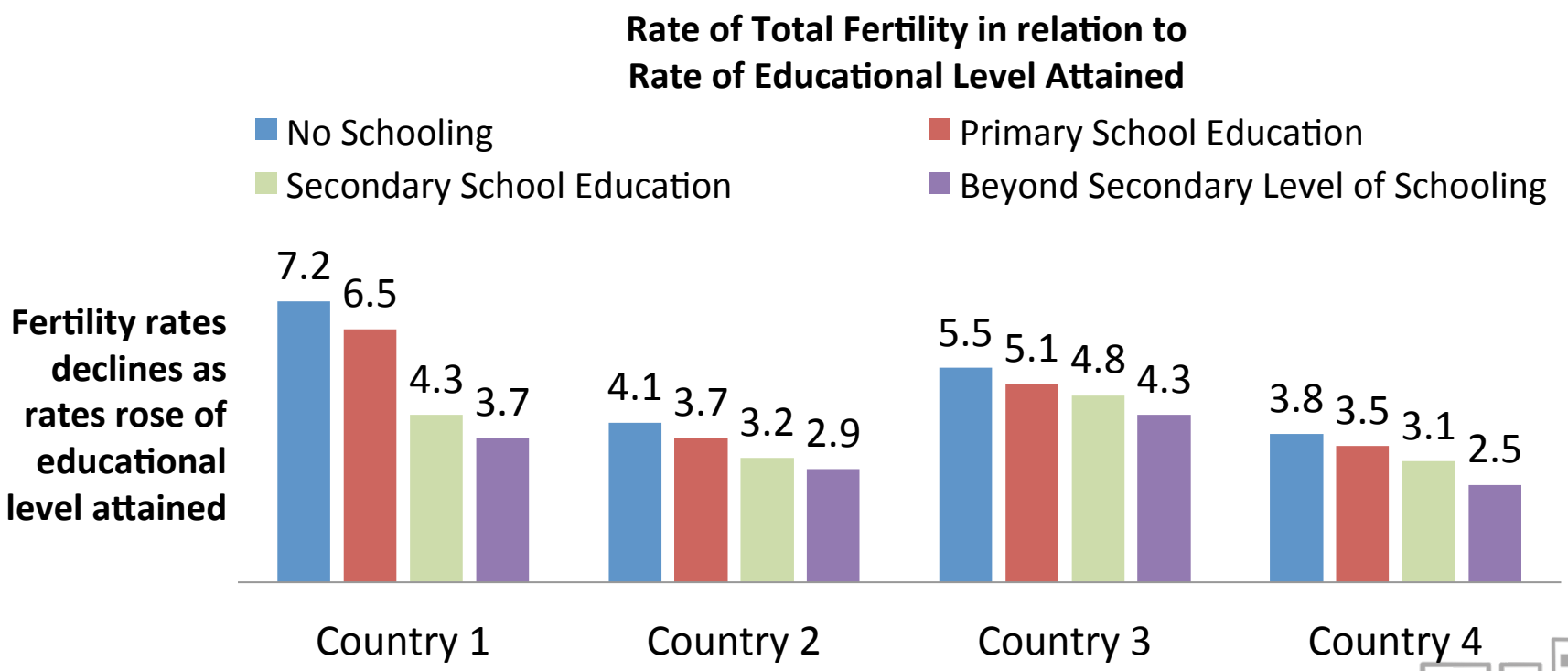
As objectives, at the end of this session, participants will be able to:

- Demonstrate an understanding of the following terms: patterns, trends, comparisons, and outliers (15 minutes)

MLE

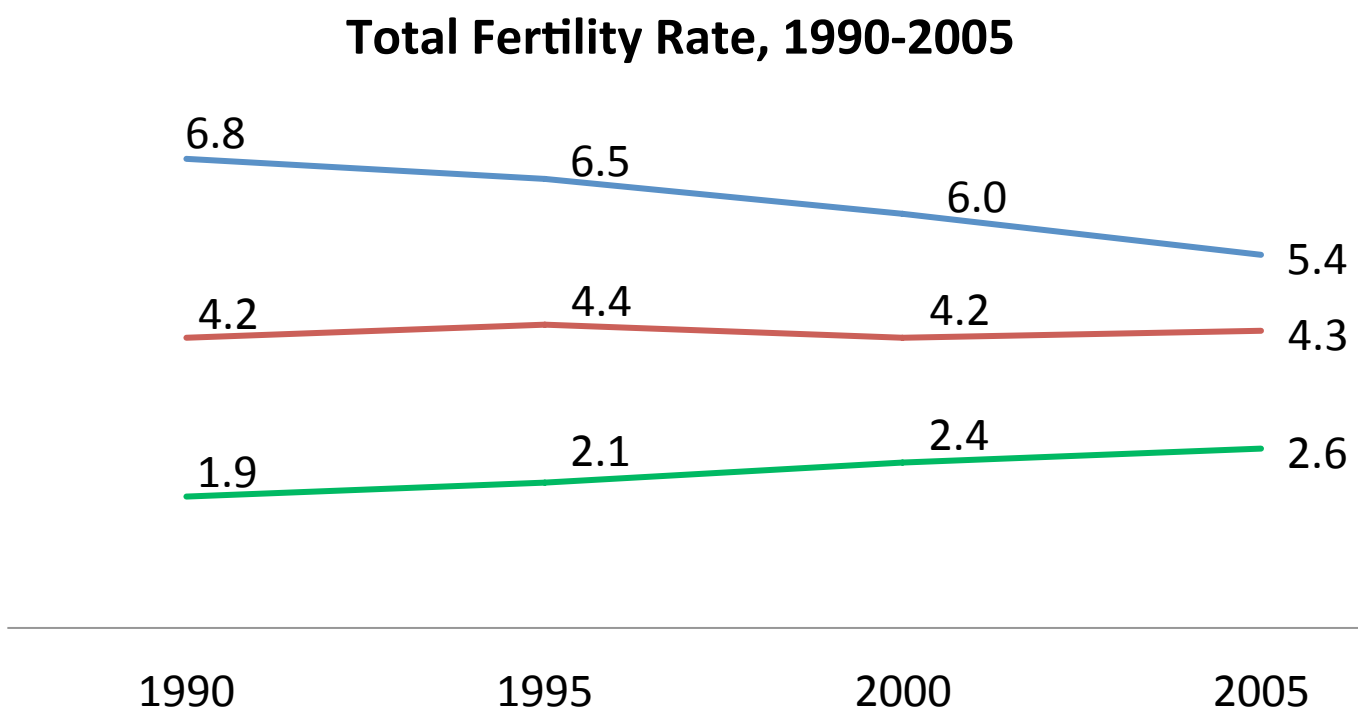
Definitions Related to Data Visualization (1)

<u>Pattern</u> : Analysis of data points to identify a regular and repeated relationship in their distribution



Definitions Related to Data Visualization (2)

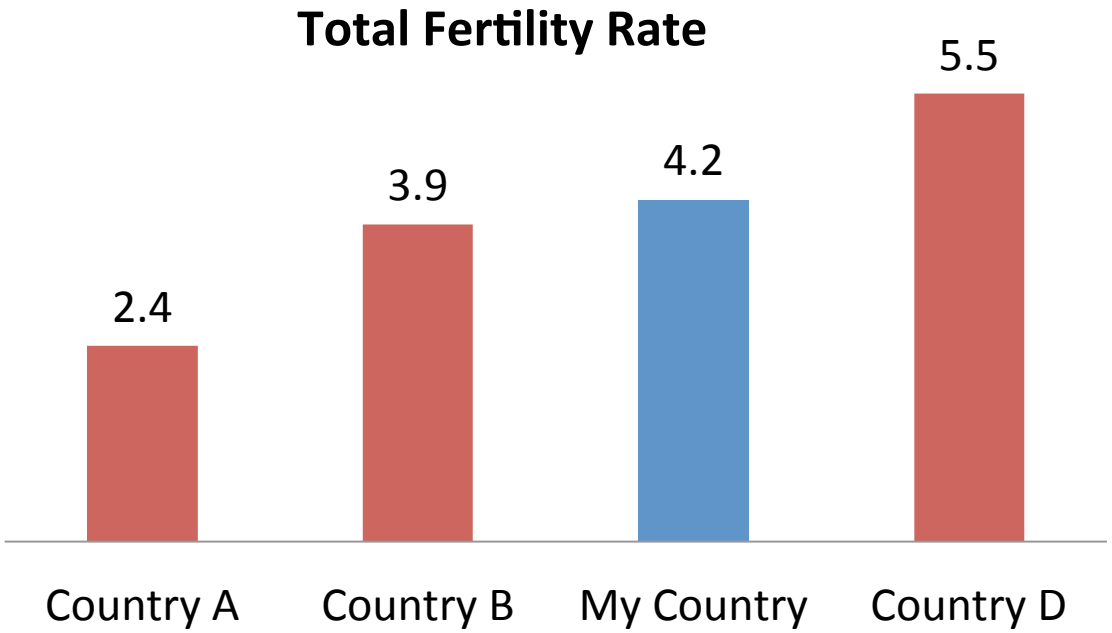
<u>Trend</u> : Analysis of data points to identify their distribution as following a certain pattern over the course of time



Definitions Related to Data Visualization (3)

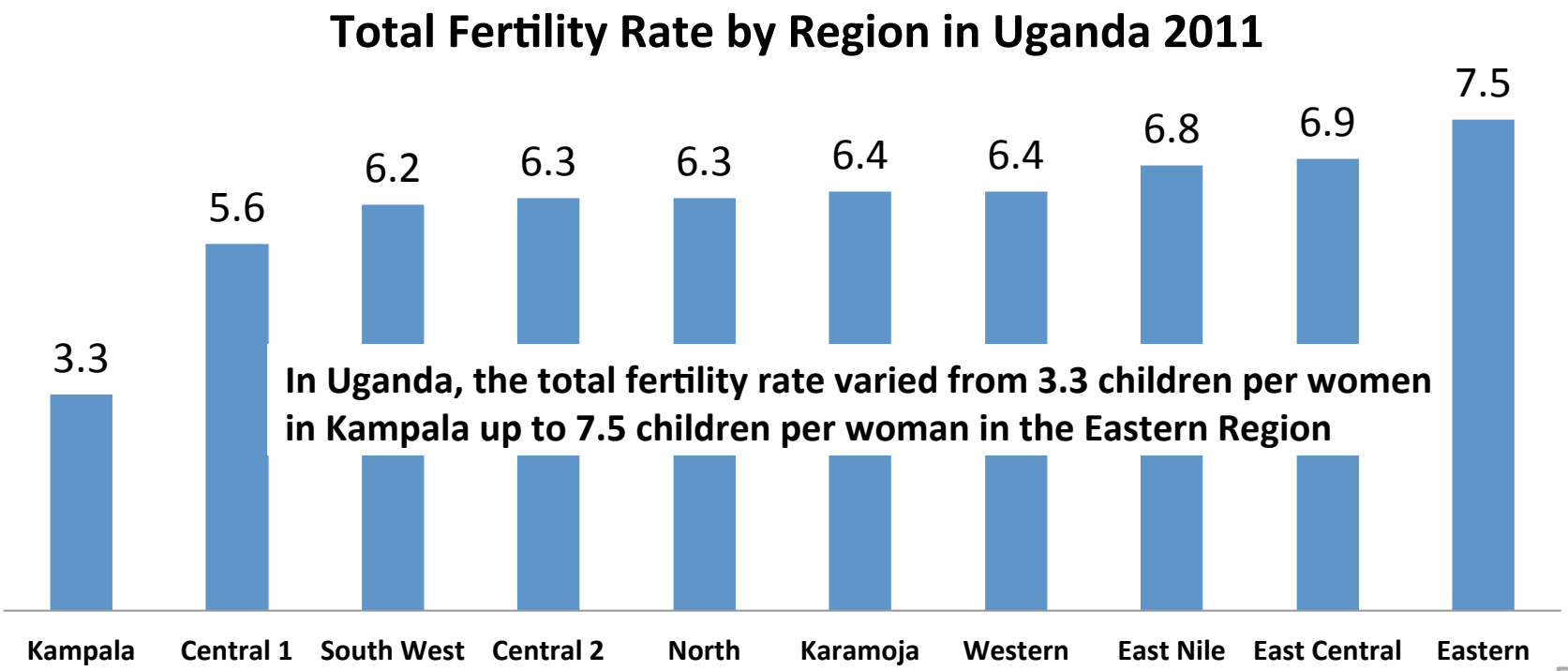
<u>Comparison</u> : Analysis of data points to identify similarities and differences among them

The total fertility rate of My Country is lower than that of Country D



Definitions Related to Data Visualization (4)

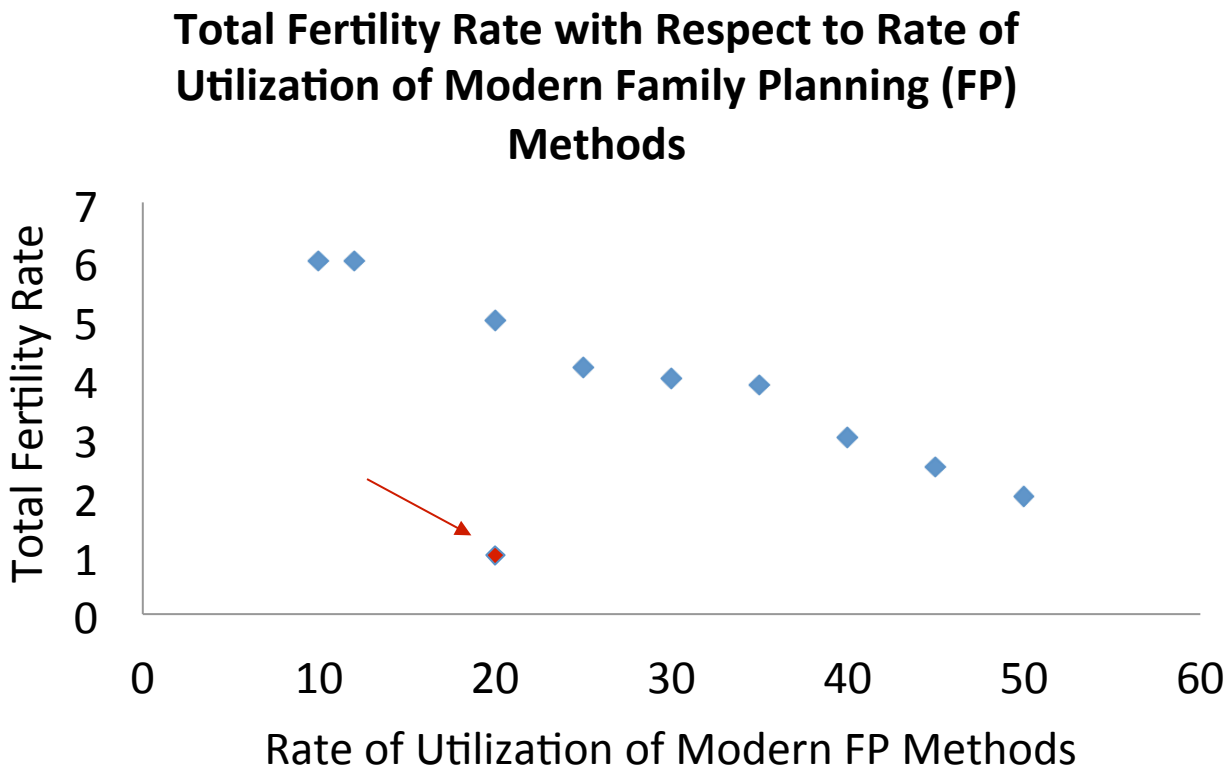
<u>Range</u> : Analysis of data points to identify the difference between the smallest and greatest value among them



MLE

Definitions Related to Data Visualization (5)

<u>Outlier</u>: Data point that is far from the others or does not follow the same pattern as the others; extreme case



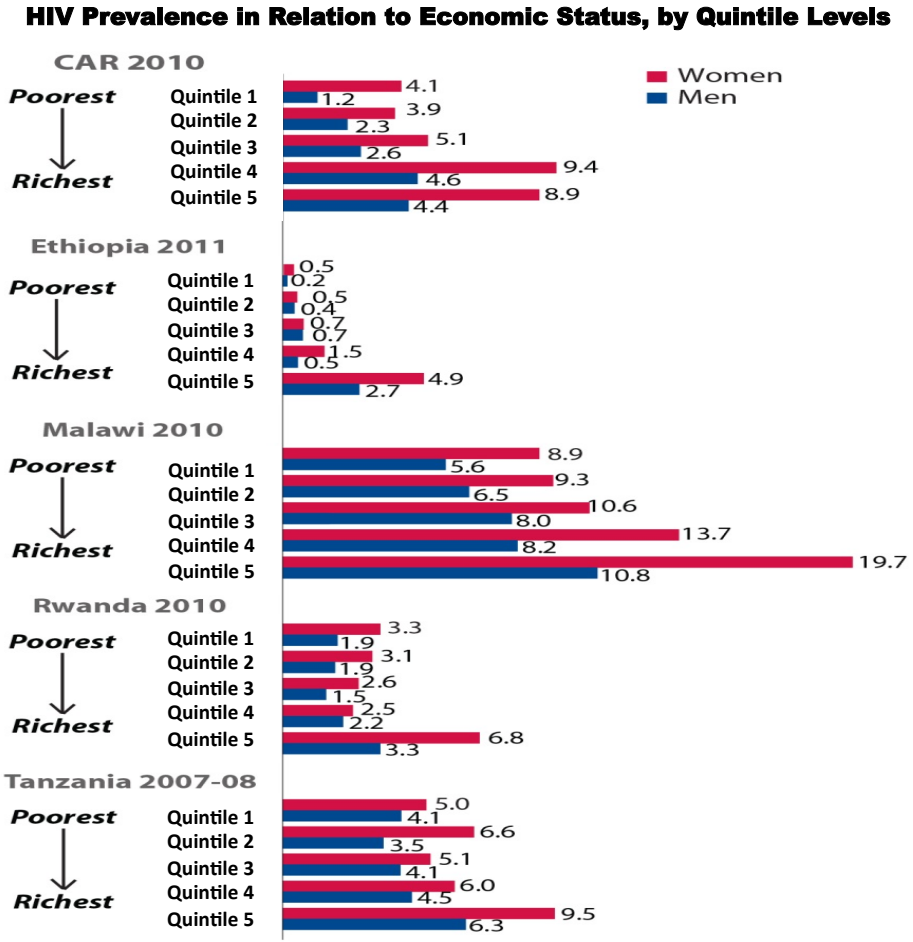
Definitions Related to Data Visualization (6)

<u>Surprise</u> : Unexpected observation from the analysis of data; unforeseen conclusion with regard to the source of the data

- Same pattern found in each country with relation to economic status level
- Comparable pattern found among all sub-Saharan African countries studied
- Surprise: these findings were in contrast to those expected to be found before the studies

=

Very Interesting Discovery

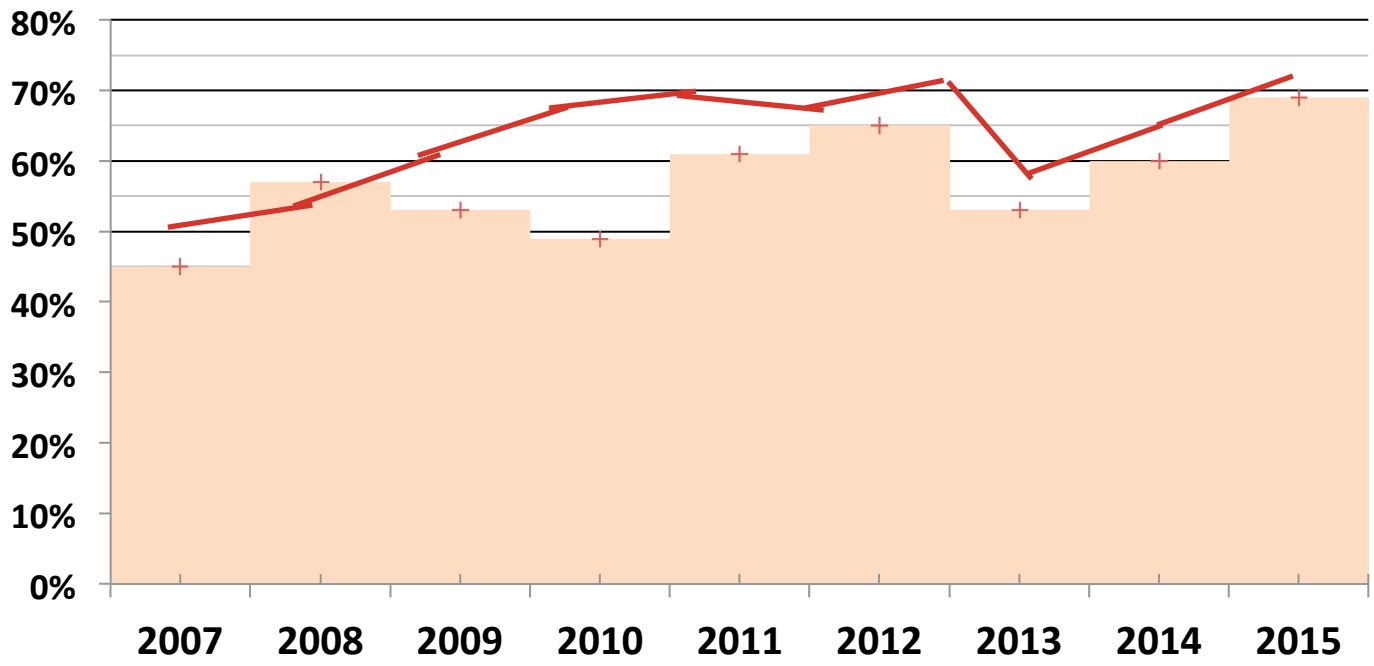


MLE

Interactive Activity #1 (15 minutes)

- Remaining in plenary, participants match one of the terms below to the conclusion to be drawn from each of the following data graphics:
  - Pattern
  - Trend
  - Comparison
  - Outlying Values

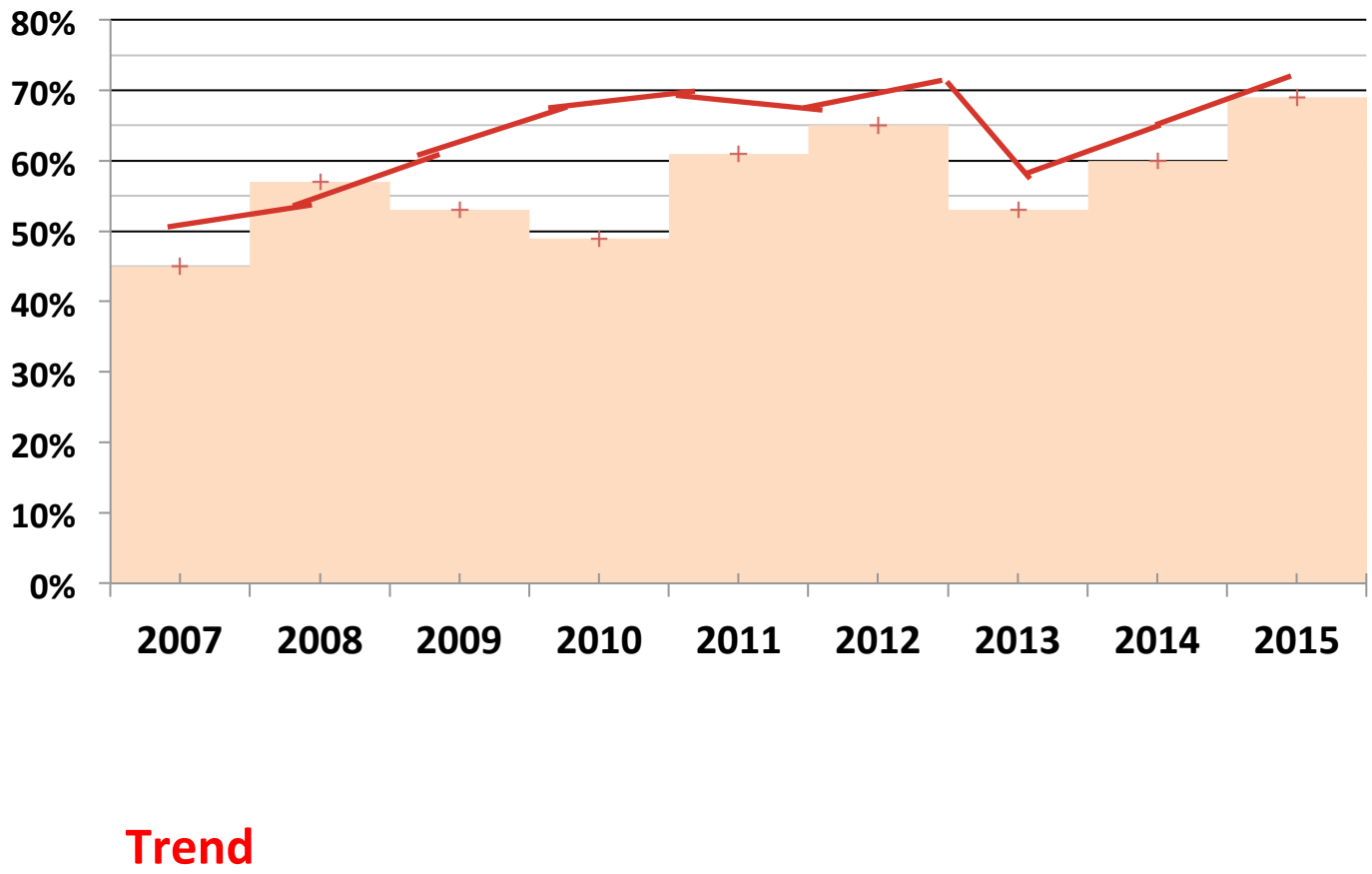
Interactive Activity #1 (15 minutes)



- Pattern?
- Trend?
- Comparison?
- Outlying Values?

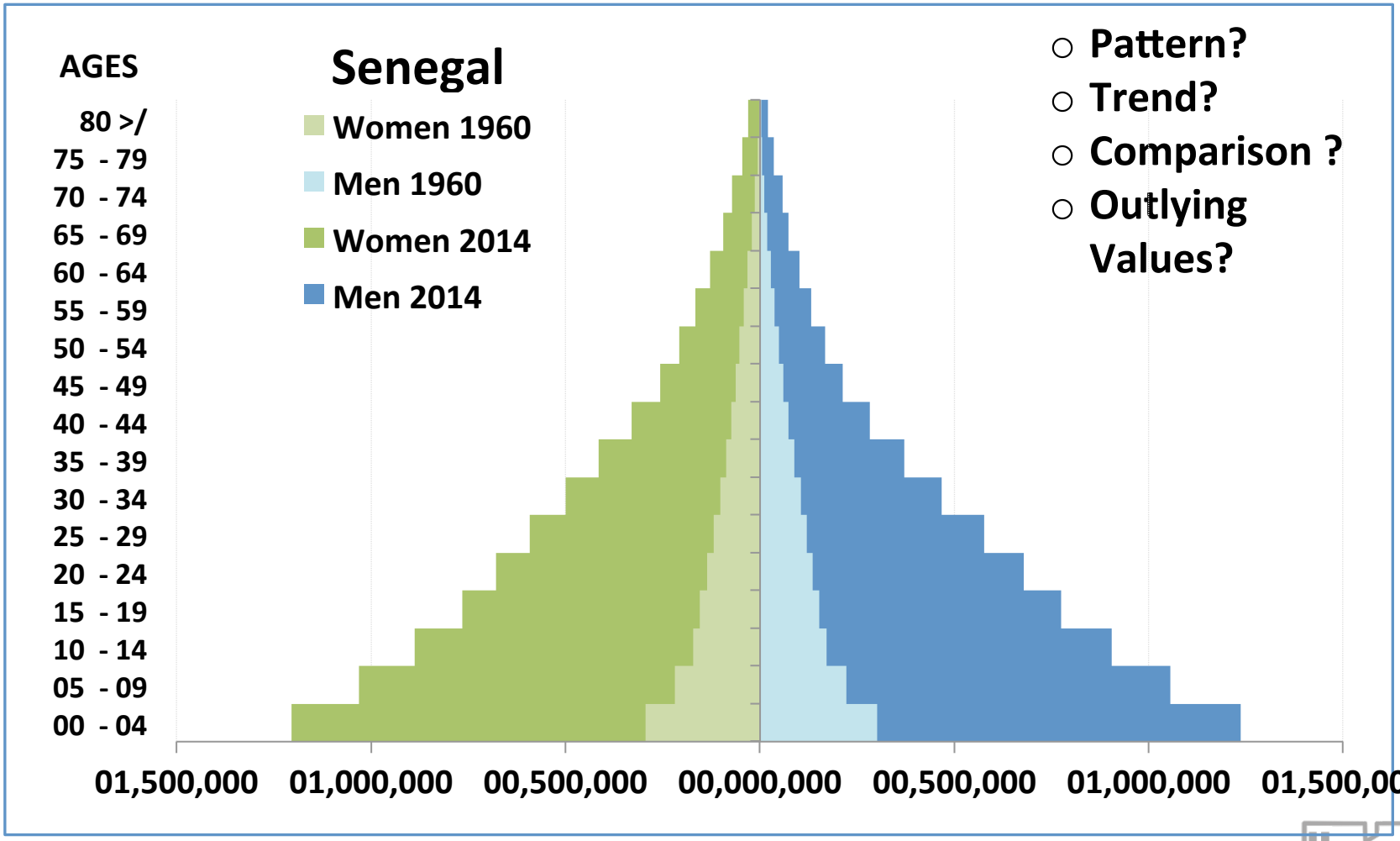
MLE

Interactive Activity #1 (15 minutes)



MLE

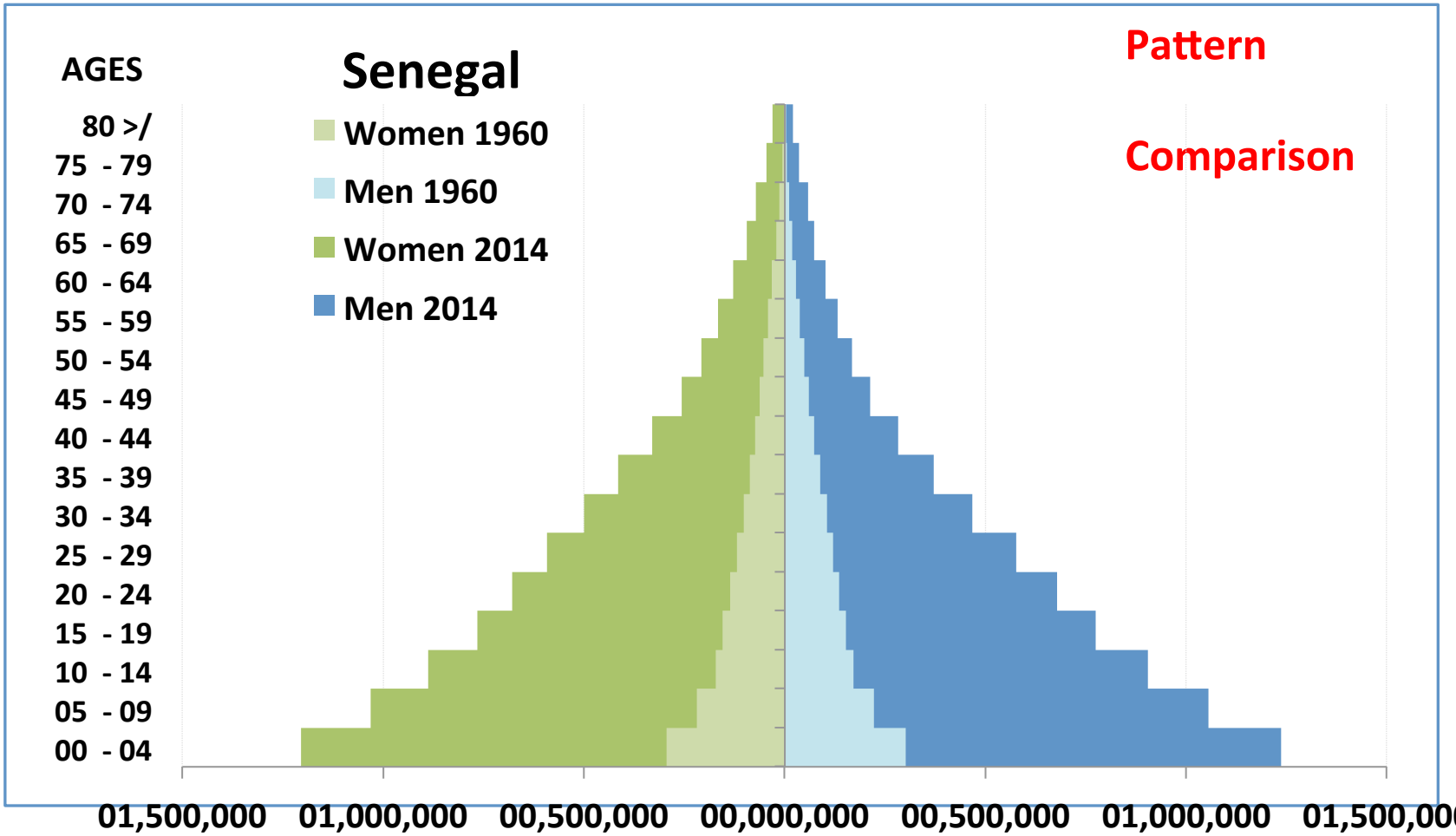
Interactive Activity #1 (continued)



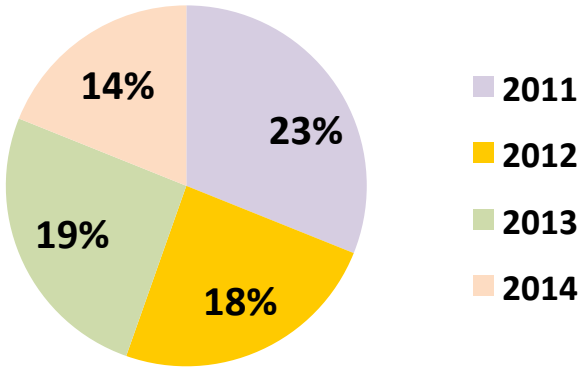
- Pattern?
- Trend?
- Comparison ?
- Outlying Values?

MLE

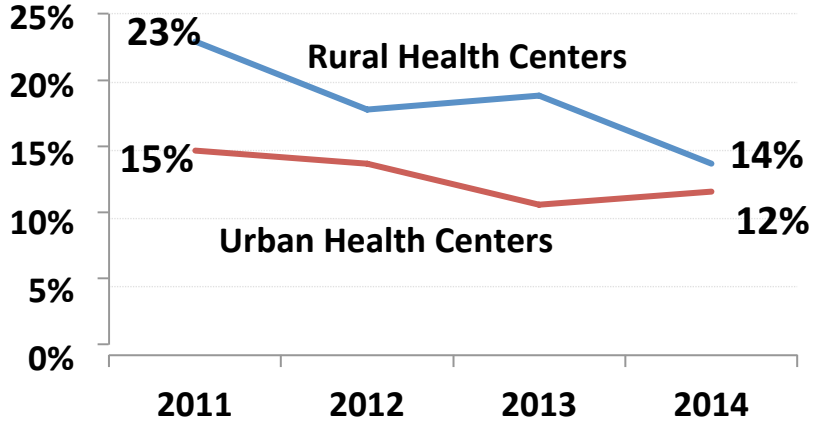
Interactive Activity #1 (continued)



Interactive Activity #1 (continued)



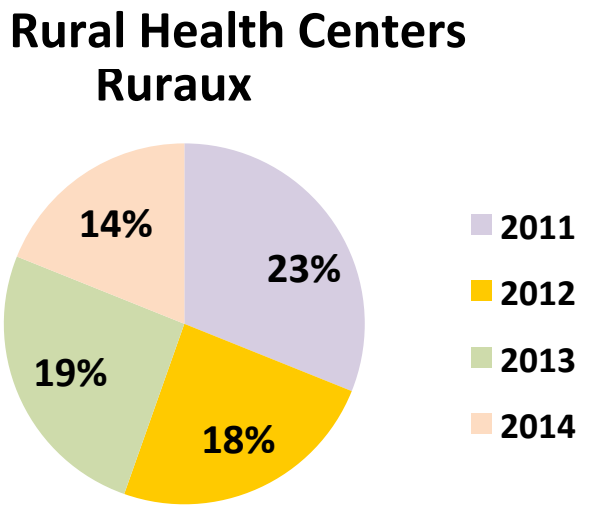
- Pattern?
- Trend?
- Comparison?
- Outlying Values?



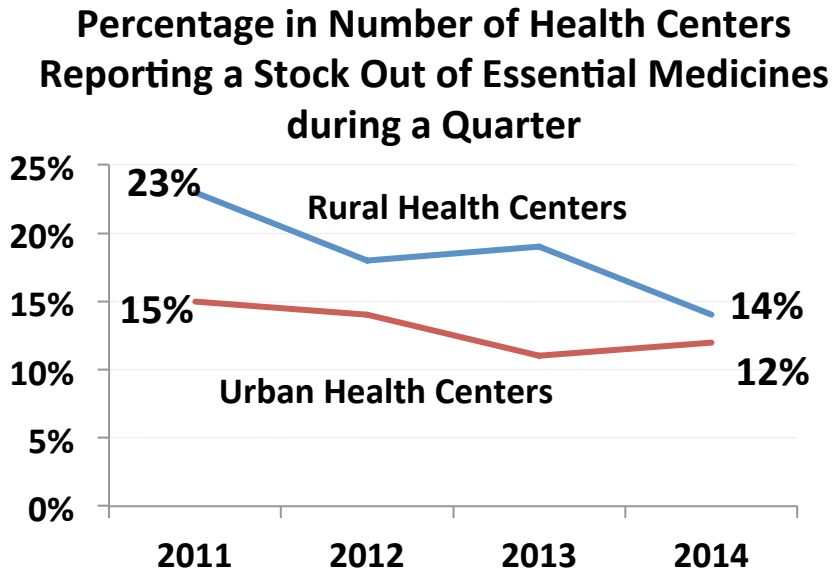
- Pattern?
- Trend?
- Comparison?
- Outlying Values?

MLE

Interactive Activity #1 (continued)



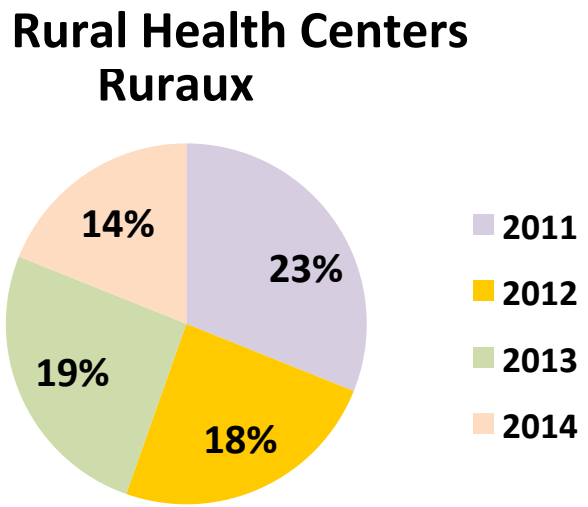
Comparison



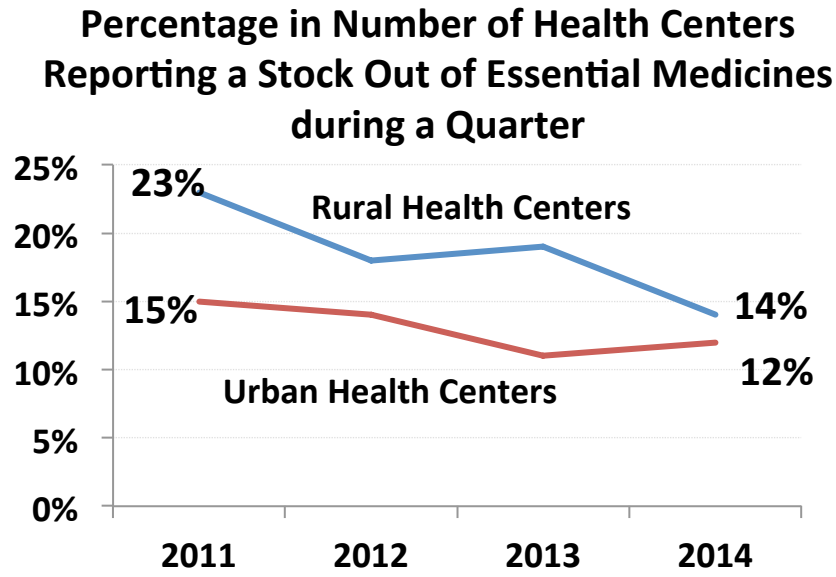
- Pattern?
- Trend?
- Comparison?
- Outlying Values?

MLE

Interactive Activity #1 (continued)



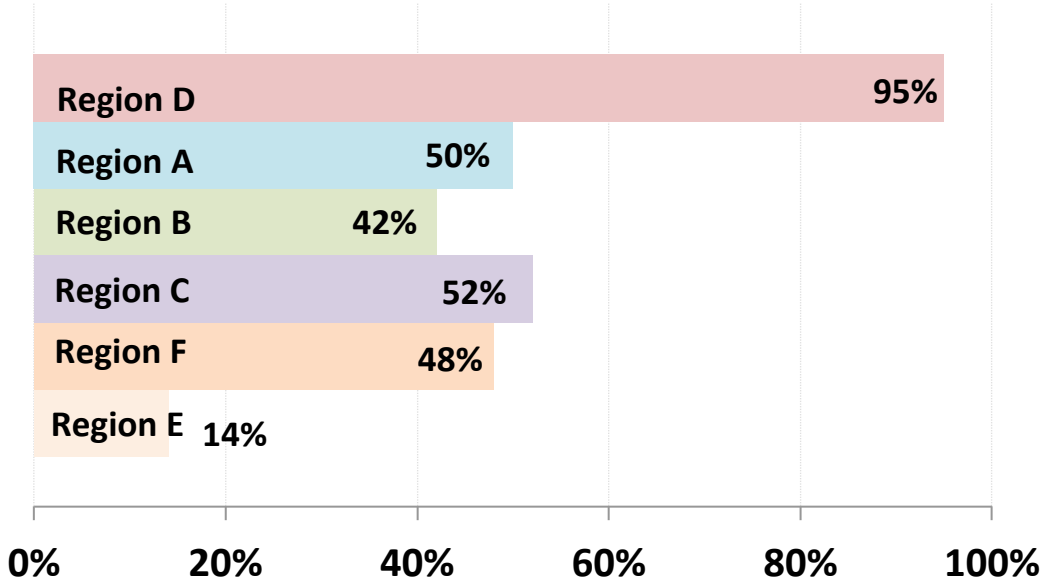
Comparison



Trend Comparison

MLE

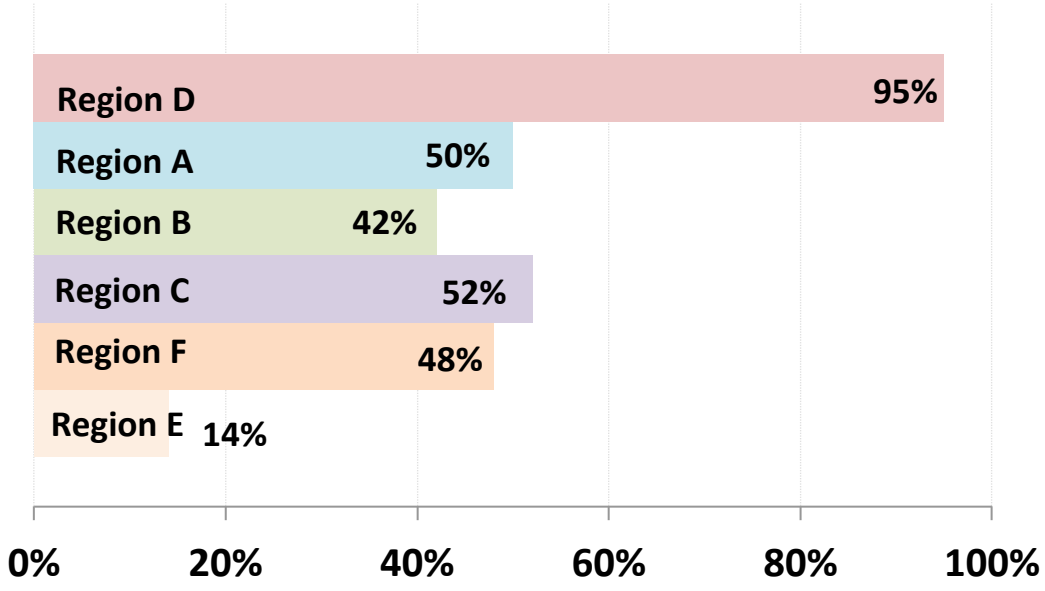
Interactive Activity #1 (continued)



- Pattern?
- Trend?
- Comparison?
- Outlying Values?

MLE

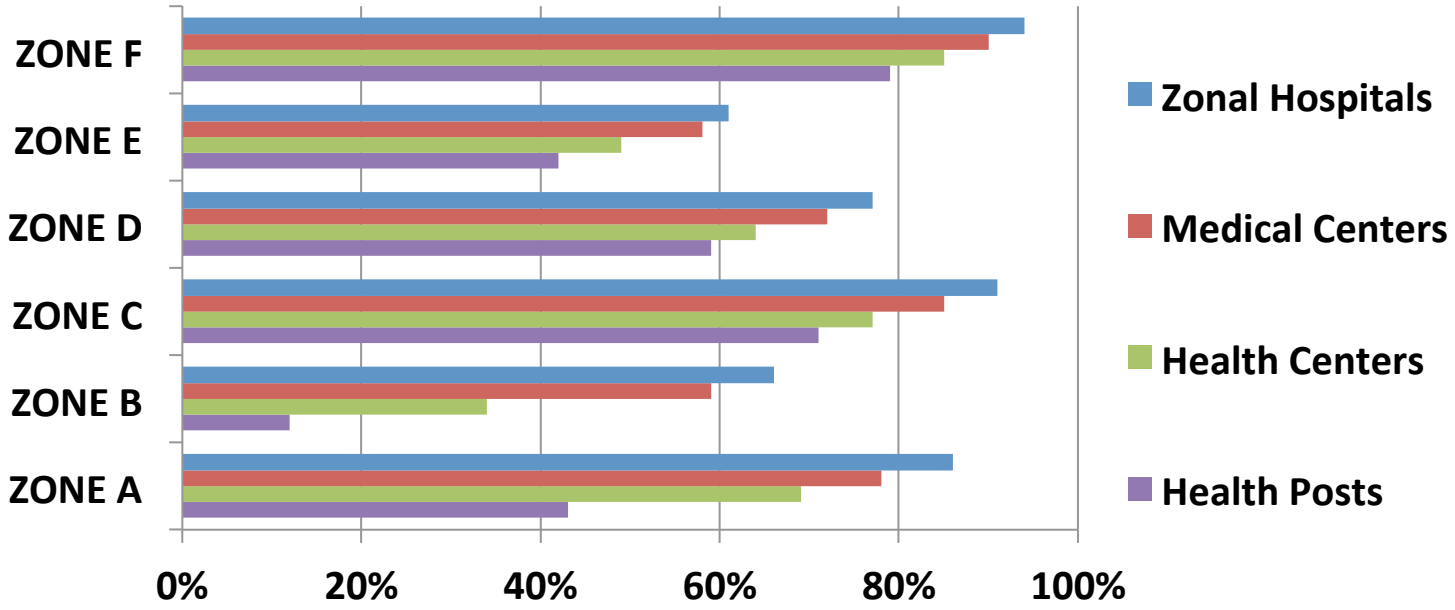
Interactive Activity #1 (continued)



Comparison
Outlying Values

MLE

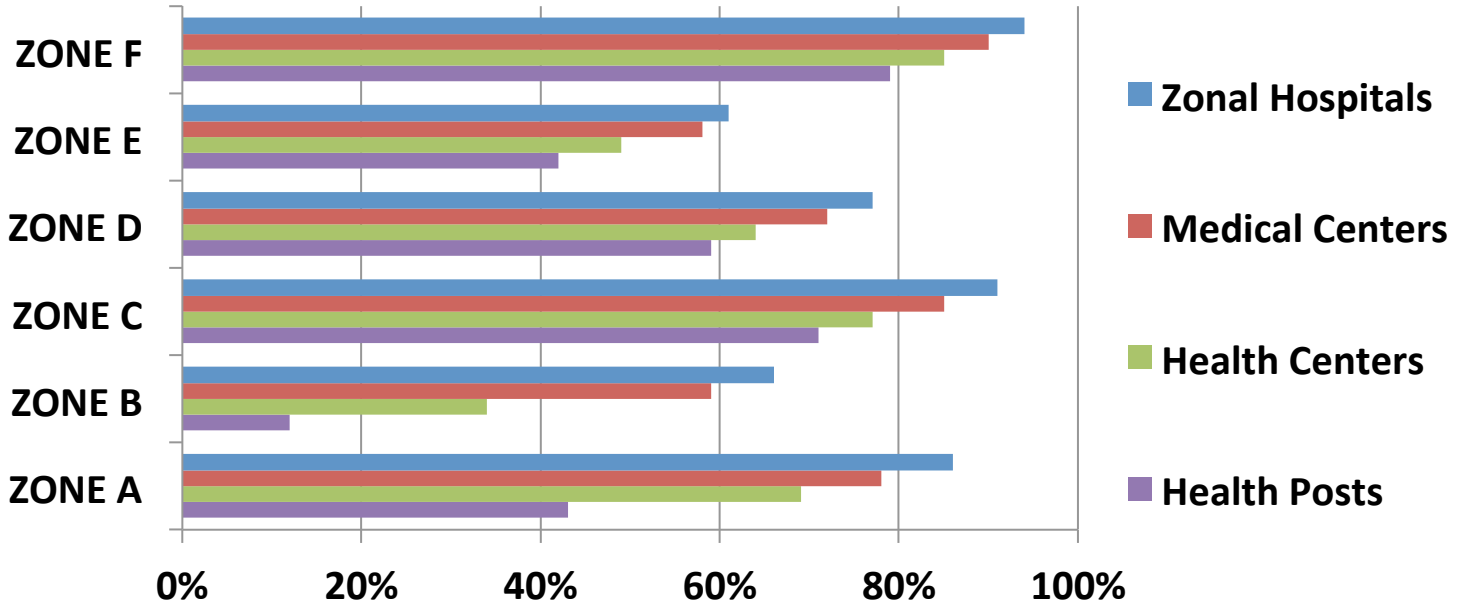
Interactive Activity #1 (continued)



- Pattern?
- Trend?
- Comparison?
- Outlying Values?

MLE

Interactive Activity #1 (continued)

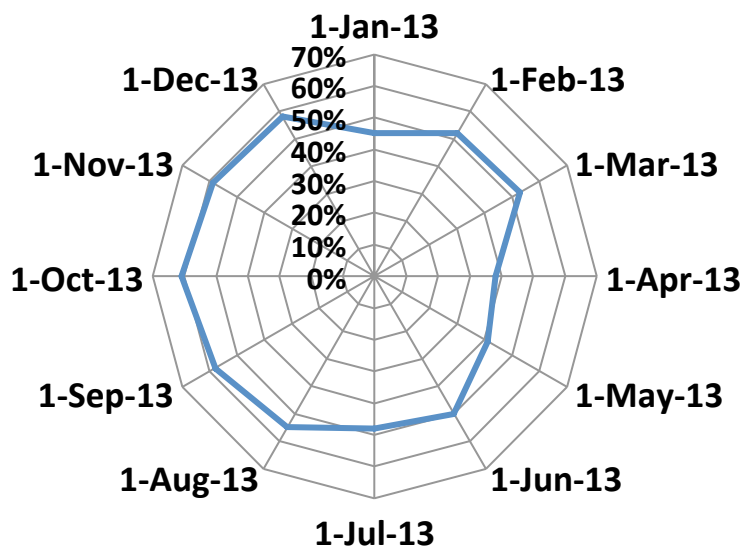


Pattern

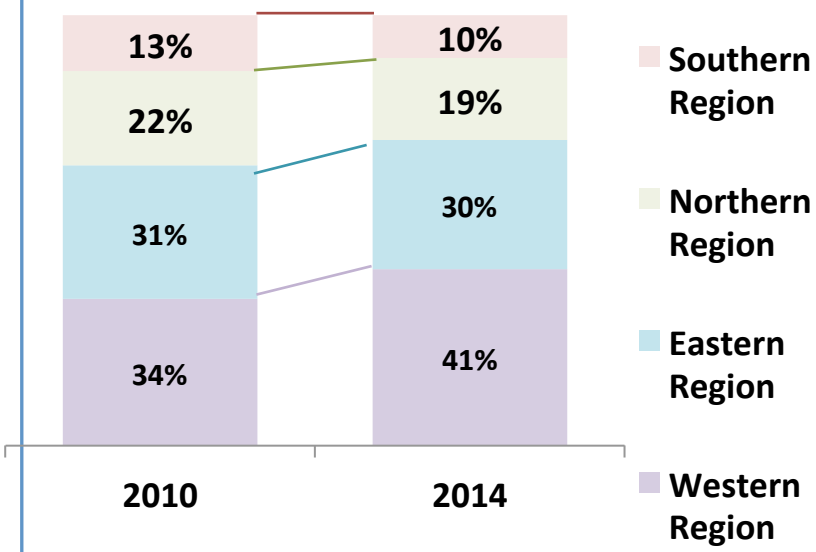
Comparison

MLE

Interactive Activity #1 (continued)



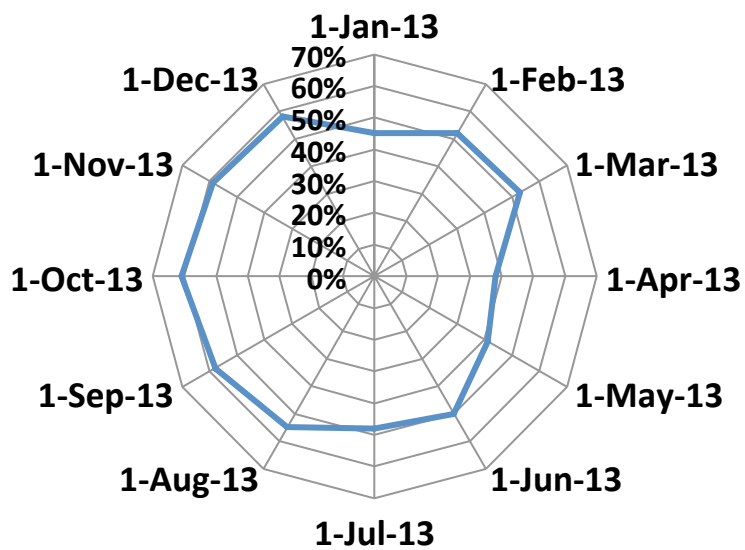
- Pattern?
- Trend?
- Comparison?
- Outlying Values?



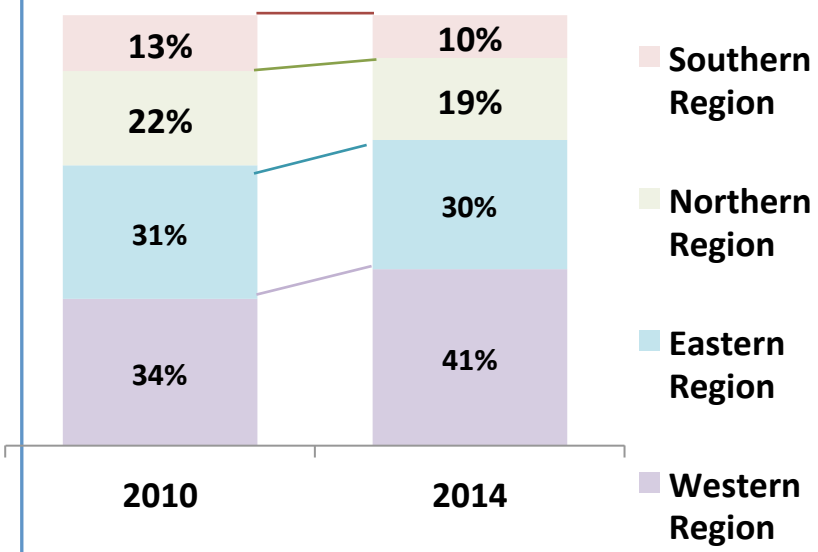
- Pattern?
- Trend?
- Comparison?
- Outlying Values?

MLE

Interactive Activity #1 (continued)

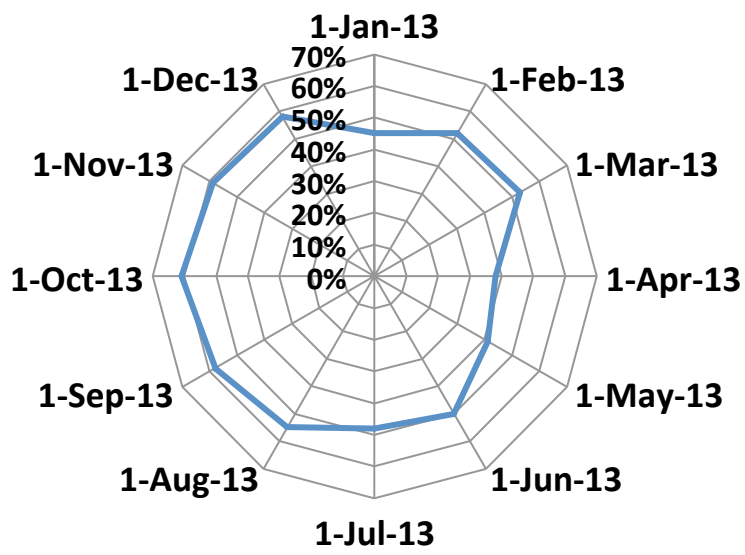


Trend

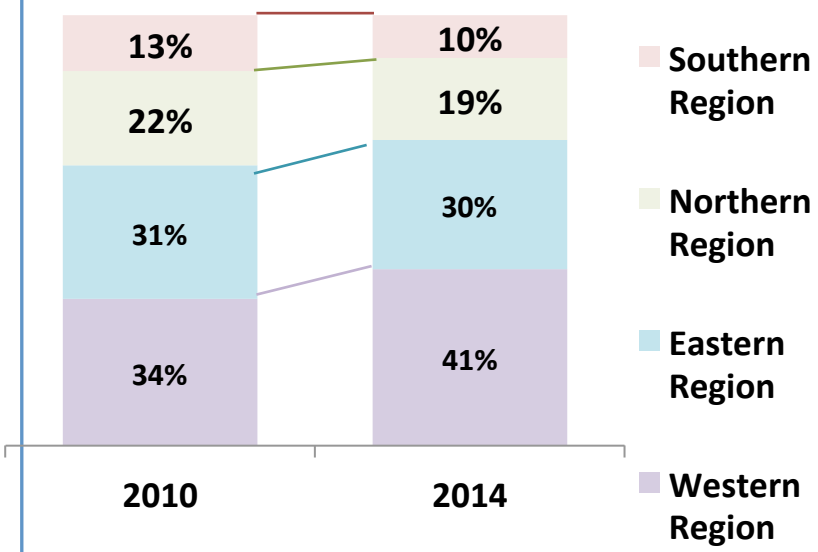


- Pattern?
- Trend?
- Comparison?
- Outlying Values?

Interactive Activity #1 (continued)



Trend



Comparison

MLE

Thank You!

Lunch Break (60 minutes)

MLE

Training in Data Visualization

Day 1, Afternoon –

MLE

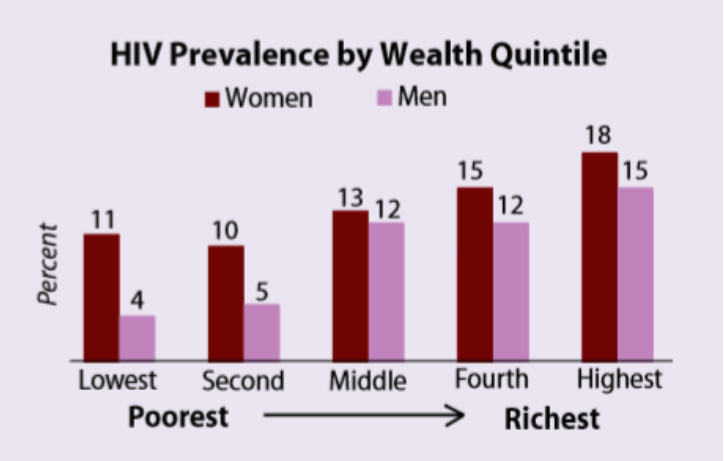
Identification of a Key Message Based on the Data: HIV and Economic Status

Table 12.4 HIV prevalence by socioeconomic characteristics

Percentage HIV positive among women and men age 15-49 who were tested, by background characteristics, Malawi 2004

Background characteristic	Women: Percentage HIV positive	Women: Number	Men: Percentage HIV positive	Men: Number	Total: Percentage HIV positive	Total: Number
Residence						
Urban	18.0	410	16.3	462	17.1	872
Rural	12.5	2,276	8.8	2,003	10.8	4,279
Region						
Northern	10.4	403	5.4	348	8.1	751
Central	6.6	1,032	6.4	994	6.5	2,026
Southern	19.8	1,251	15.1	1,122	17.6	2,373
District						
Blantyre	22.5	211	22.1	247	22.3	457
Kasungu	5.5	116	2.8	116	4.1	232
Machinga	14.9	99	8.2	86	11.8	185
Mangochi	21.4	136	19.9	108	20.8	244
Mzimba	6.4	178	3.9	157	5.2	336
Salima	9.5	74	8.2	59	8.9	133
Thyolo	23.1	145	18.6	126	21.0	271
Zomba	24.6	134	10.5	124	17.8	258
Lilongwe	a	352	a	398	a	750
Mulanje	23.3	117	15.1	86	19.8	202
Other districts	12.6	1,124	8.7	958	10.8	2,082
Education						
No education	13.6	667	9.2	295	12.3	962
Primary 1-4	12.3	690	6.5	550	9.7	1,240
Primary 5-8	13.2	958	10.8	957	12.0	1,916
Secondary+	15.1	370	12.9	662	13.7	1,032
Employment						
Currently working	14.6	1,545	13.4	1,360	14.0	2,904
Not currently working	[illegible]	1,141	[illegible]	1,105	9.0	2,246
Wealth quintile						
Lowest	10.9	455	4.4	300	8.3	755
Second	10.3	546	4.6	487	7.6	1,033
Middle	12.7	581	12.1	568	12.4	1,149
Fourth	14.6	595	11.7	564	13.2	1,159
Highest	18.0	508	14.9	546	16.4	1,054
Ethnicity						
Chewa	[illegible]	915	[illegible]	796	7.1	1,711
Tumbuka	8.8	280	5.1	231	7.1	511
Lomwe	20.0	450	17.9	433	19.0	884
Tonga	15.9	53	(9.2)	49	12.6	102
Yao	17.5	337	12.4	300	15.1	637
Sena	12.4	126	10.3	115	11.4	241
Nkonde	(19.1)	30	4.9	51	10.2	82
Ngoni	14.5	288	14.5	299	14.5	587
Other	17.3	204	9.2	190	13.4	394
Religion						
Catholic	13.8	634	10.5	552	12.3	1,185
CCAP	9.7	481	8.7	464	9.2	945
Anglican	17.7	56	(5.6)	47	12.1	103
Seventh Day Adventist/Baptist	12.1	158	16.5	175	14.4	334
Other Christian	13.8	1,025	10.0	890	12.0	1,915
Muslim	17.0	309	11.3	251	14.4	560
No religion	*	20	0.0	75	0.7	95
Total	13.3	2,686	10.2	2,465	11.8	5,150

Malawi 2004 Demographic and Health Survey



MLE

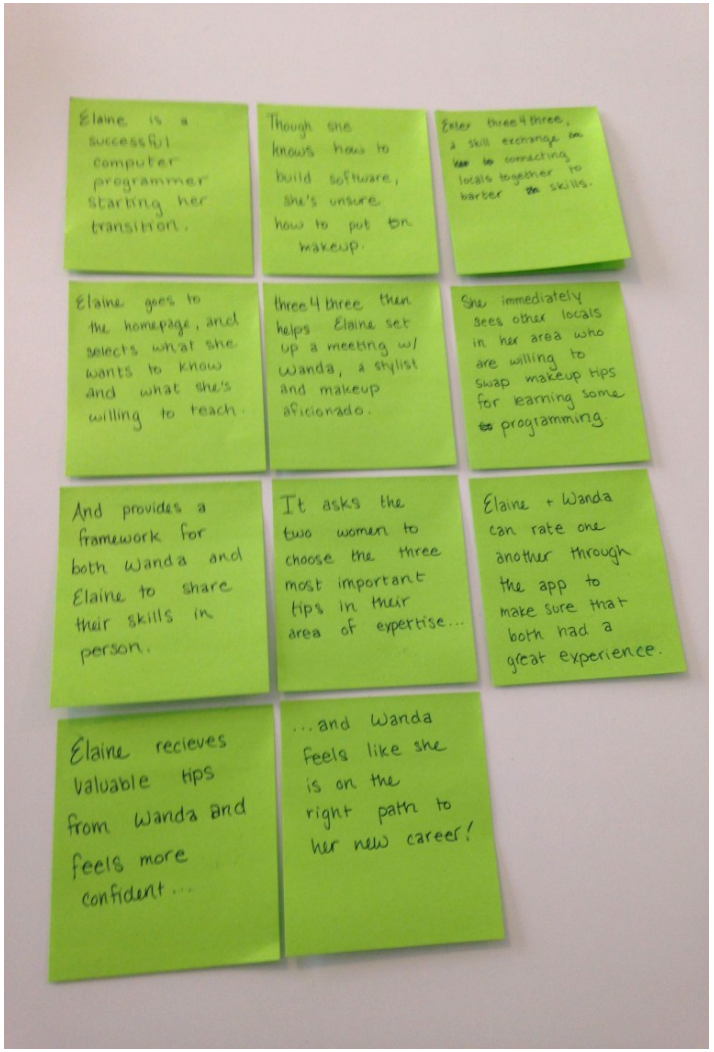
Identification of a Key Message to Transmit, after it been Formulated from Analysis of the Data

As objectives, at the end of this session, participants will be able to

- Demonstrate an understanding of the method for storyboarding (40 minutes)

MLE

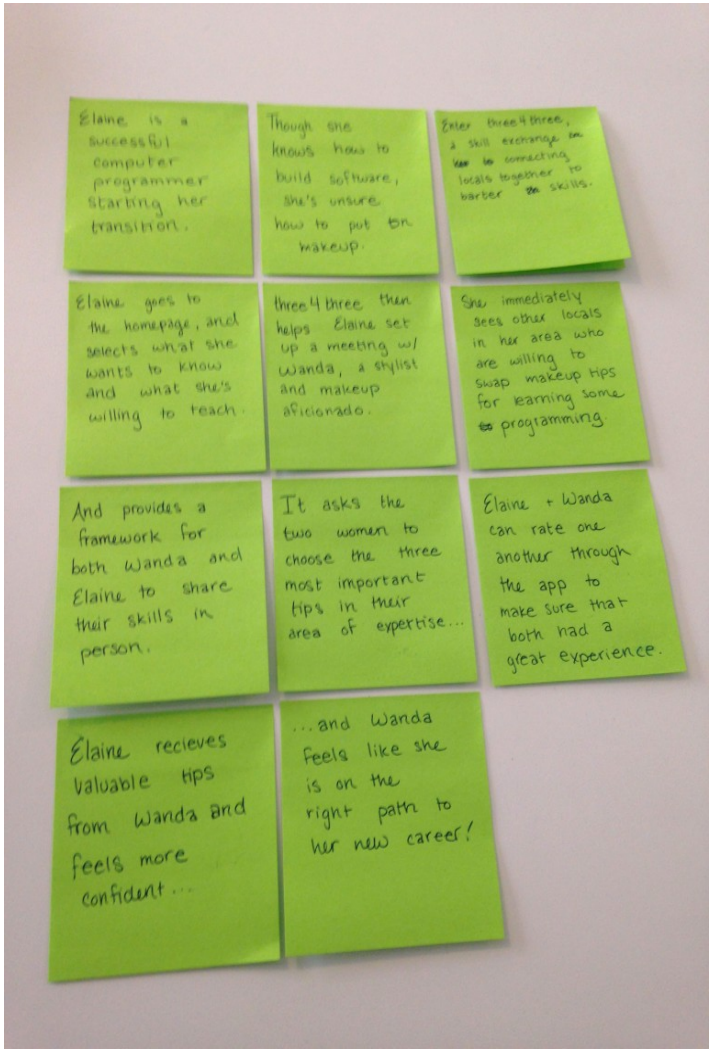
Identification of a Key Message by Means of Storyboarding: Explanation



- Approach for collecting and ordering ideas for formulating a message that is based on the analysis of the data
- Collaborative process for gaining a consensus from all those involved in transmitting the message
- Strategy for evaluating the available information in all its variety by mapping meaningful elements into a series to form a narrative flow

MLE

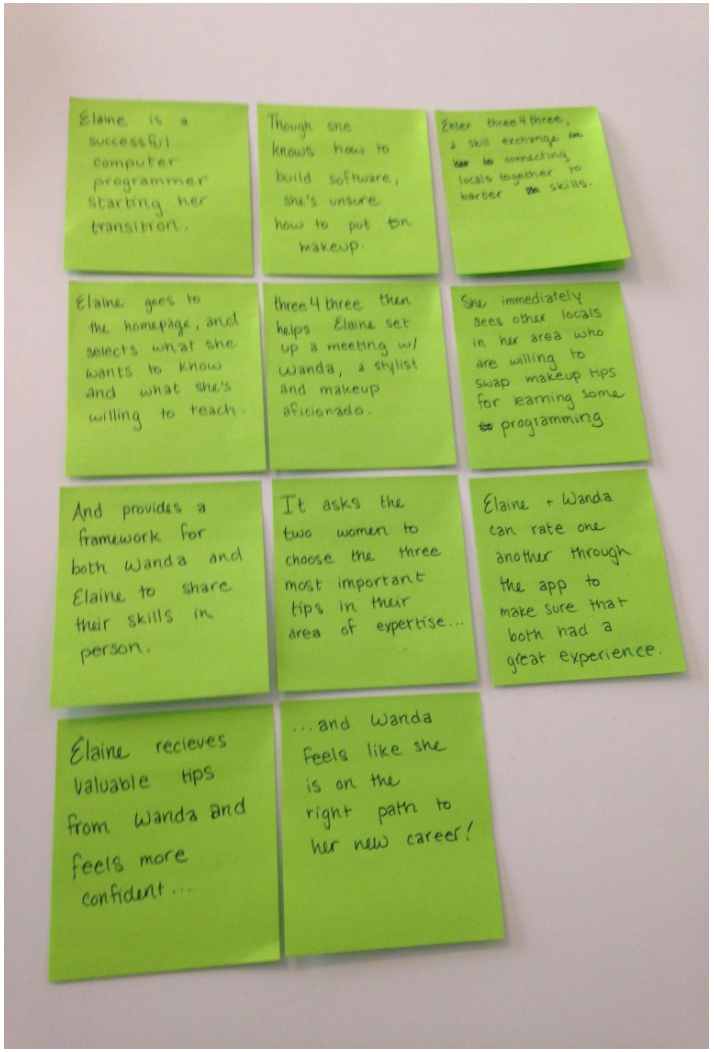
Identification of a Key Message by Means of Storyboarding: Procedure (1)



- Write on pieces of paper the essential items of information that have been identified from an analysis of the data for planned dissemination to targeted audiences
- Attach these pieces of paper on a wall and arrange them in a way that reveals the narrative ties between them

MLE

Identification of a Key Message by Means of Storyboarding: Procedure (2)



- Rearrange these pieces of paper if there are more that are added or if the message needs to be refined or if certain elements must be prioritized
- Formulate a single sentence that can serve as an overall and definitive summary
- Document or photograph the final arrangement

MLE

Interactive Activity #2 (25 minutes)

- Separate yourselves into pairs/groups in order to undertake the process of storyboarding:
  - Write on Post-Its the different conclusions that each pair/ group was able to draw from the following data graphic
  - Arrange the order of these Post-Its in order to develop a narrative flow by which the conclusions are interrelated to each other
  - Arrive at an overall and definitive statement that can be reduced to a single sentence
  - Present this summary statement to the plenary group for comments

MLE

Interactive Activity #2 (continued)

(25 minutes)

	Frequency of Attendance at Births in the Eastern Region during 2014
Doctors	**42.0 %**
Traditional Midwives	**41.0 %**
Trained Midwife / Nurse / Female Health Agent	**10.2 %**
Family Member	**6.2 %**
No One	**.2 %**
Lack of Information	**.4 %**

MLE

Identification of a Key Message to be Told that has been Formulated from an Analysis of the Data

As objectives, at the end of this session, participants will be able to:

- Understand the considerations that must be held by the MLE Project MLE in Kenya regarding its dissemination of findings from its field surveys (75 minutes)

MLE

Interactive Activity #3 (80 minutes)

- Separate yourselves into pairs/groups in order to arrive at an overall and definitive summary statement based on the analysis of a set of data tables that will be furnished to you from the MLE Project in Kenya
- Take up again the steps that have been presented to you during this training up to now:

MLE

Interactive Activity #3 (continued)

- Specify the characteristics of a particular, targeted audience and identify its informational needs
- Determine if your data table(s) reveals a pattern, a trend, a comparison, or outlying values
- Undertake the storyboarding process in order to tie together these conclusions to each other
- Formulate an overall and definitive summary statement based on meaningful findings from your analysis of the data tables
- Justify to the plenary group the effectiveness of this summary statement in reaching its target audience

MLE

Distribution (in percentage) of women at KENYA MLE Project sites in a relationship, according to their current usage of a contraceptive method at the time of the baseline survey (2010) and end-line survey (2014)

% SURVEYED WOMEN in a RELATIONSHIP Baseline Survey Results End-Line Survey Results	USING ANY METHOD	Any Modern Method	Any Traditional Method	*Non-User*
NAIROBI	47.8 61.6	43.6 54.8	4.2 6.8	*52.2* *38.4*
MOMBASA	33.8 47.1	29.3 43.8	4.5 3.3	*66.3* *52.9*
KISUMU	47.9 62.5	44.4 58.7	3.5 3.8	*52.1* *37.5*
MACHAKOS	53.3 64.6	45.3 57.8	8.0 6.8	*46.8* *35.4*
KAKAMEGA	48.5 55.4	46.1 53.8	2.4 1.6	*51.5* *44.5*

Examples of Data from the Field Studies Carried out by the MLE Project

Distribution (in percentages) of Women at MLE Kenya Project Sites According to their Knowledge (Spontaneous Knowledge or After Prompting) of Contraceptive Methods at the Time of MLE Project Baseline Survey (2010) & End-line Survey (2014)

	NAIROBI		MOMBASA		KISUMU	
	Project Baseline 2010	Project End-Line 2014	Project Baseline 2010	Project End-Line 2014	Project Baseline 2010	Project End-Line 2014
	%	%	%	%	%	%
ANY MODERN METHOD	**98.1**	**100.0**	**98.3**	**99.8**	**99.0**	**99.9**
Female Sterilization	63.4	90.8	71.1	80.0	74.8	86.0
Male Sterilization	50.0	76.5	56.7	62.8	52.7	65.6
IUD	82.8	97.1	78.0	95.4	83.3	96.4
Injectables	95.4	99.3	93.4	98.3	96.5	99.4
Implants	81.4	98.0	78.9	94.7	87.0	98.1
Daily Pills	94.9	99.4	94.4	96.8	95.0	99.4
Emergency Contraception	57.7	79.4	44.5	57.6	57.6	67.3
Male Condom	96.7	99.8	97.5	98.7	98.4	99.8
Female Condom	87.1	96.9	74.9	85.7	87.8	95.2
LAM Breastfeeding	49.7	76.7	40.2	77.3	46.8	52.6
Std Days Method, CycleBeads	79.2	65.5	78.9	46.3	78.0	45.7
***TRADITIONAL** -- Withdrawal*	*56.9*	*87.2*	*62.6*	*69.4*	*56.2*	*69.7*

Examples of Data from the Field Studies Carried out by the MLE Project

Distribution (in percentages) of Women at MLE Kenya Project Sites According to their Knowledge (Spontaneous Knowledge or After Prompting) of Contraceptive Methods at the Time of MLE Project Baseline Survey (2010) & End-line Survey (2014)

	MACHAKOS		KAKAMEGA	
	Project Baseline 2010	Project End-Line 2014	Project Baseline 2010	Project End-Line 2014
	%	%	%	%
ANY MODERN METHOD	**98.9**	**100.0**	**98.4**	**100.0**
Female Sterilization	83.8	89.4	76.6	93.6
Male Sterilization	58.3	57.2	65.7	74.0
IUD	85.0	93.8	87.3	97.8
Injectables	95.8	98.3	95.9	99.8
Implants	86.3	96.3	84.0	98.6
Daily Pills	95.5	98.0	94.2	99.8
Emergency Contraception	52.8	51.2	51.9	63.7
Male Condom	98.0	99.7	97.4	99.9
Female Condom	80.0	86.6	90.5	95.2
LAM Breastfeeding	37.7	41.2	54.5	63.2
Std Days Method, CycleBeads	87.5	37.0	81.6	34.6
***TRADITIONAL** -- Withdrawal*	*62.6*	*67.4*	*63.1*	*73.5*

Coffee / Tea Break (15 minutes)

Reflections

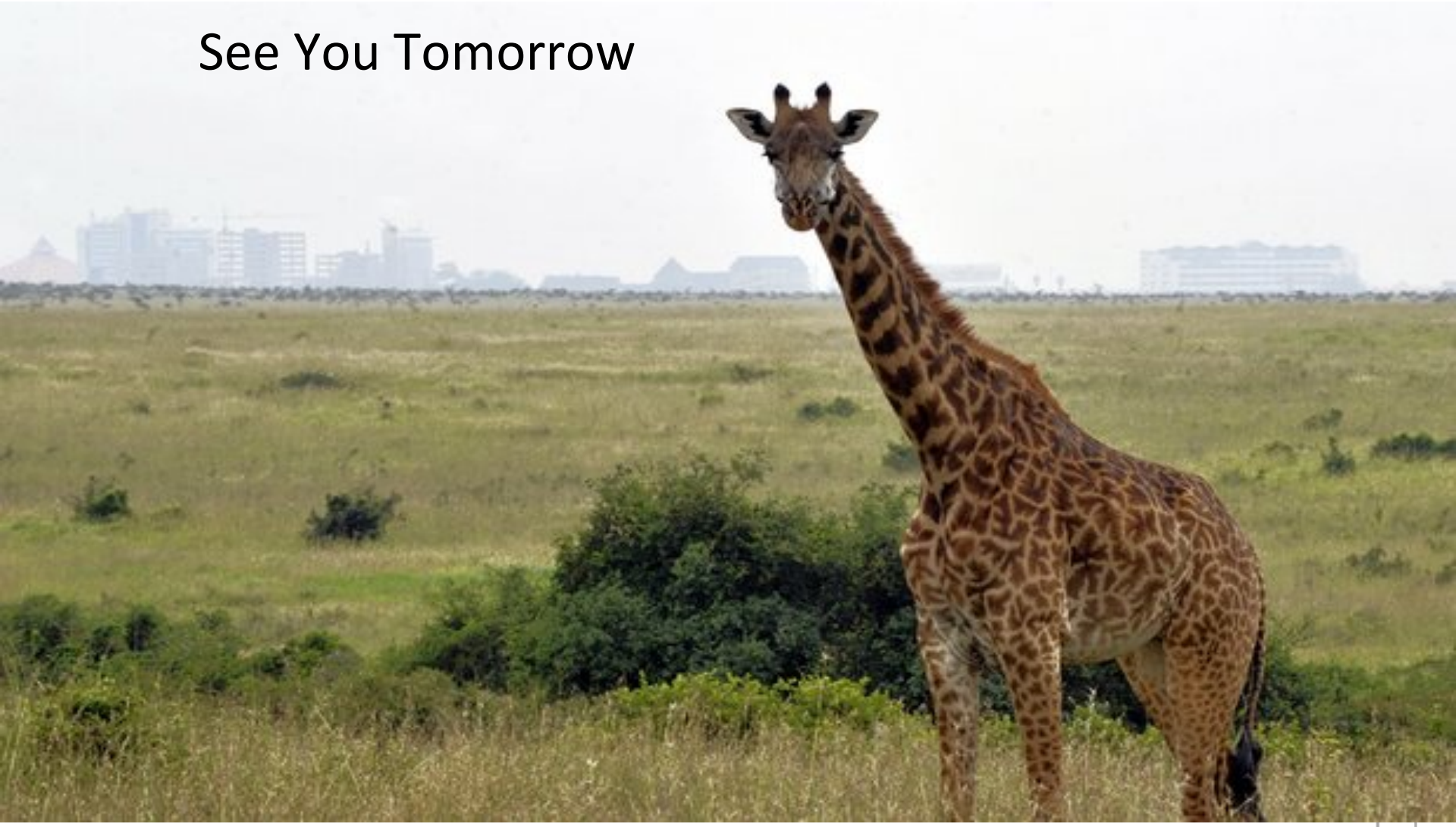
- While in plenary, provide a summary of the topics covered during the course of today (20 minutes)

- Separate yourselves into pairs/groups in order to evaluate the quality of instruction of the morning's and afternoon's session by means of the graphic method of assessment that you had earlier developed (10 minutes)

MLE

Thank You!

See You Tomorrow



MLE

Training in Data Visualization

Day 2, Morning –

MLE

Retrospective and Prospective

- Discuss within the plenary group a summary of topics that were covered yesterday (5 minutes)

- Presentation by each pair of its evaluation of the quality of instruction from the previous morning's and afternoon's session by means of its graphic method of assessment (10 minutes)

- Summarization of topics to be covered today (5 minutes)

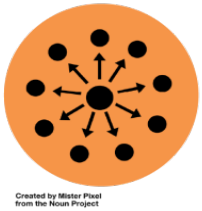
MLE

Process: Developing a Data Visualization

1. Identify your audience & context
- Who are the stakeholders?
- What do they need to know?

2. Find the story in your data
- What information do you have?
- Are there identifiable patterns, trends, surprises, relationships, successes or failures?

3. Build your Visualization
- Sketch first!
- Choose the right type of chart for your data
- Simplify
- Use pre-attentive attributes to highlight your story



4. Disseminate, share, and use!
- How will this visualization be used?
- What is the ideal format?
- Good data visualization ensures that data are read, understood, and used for evidence-based decision-making.

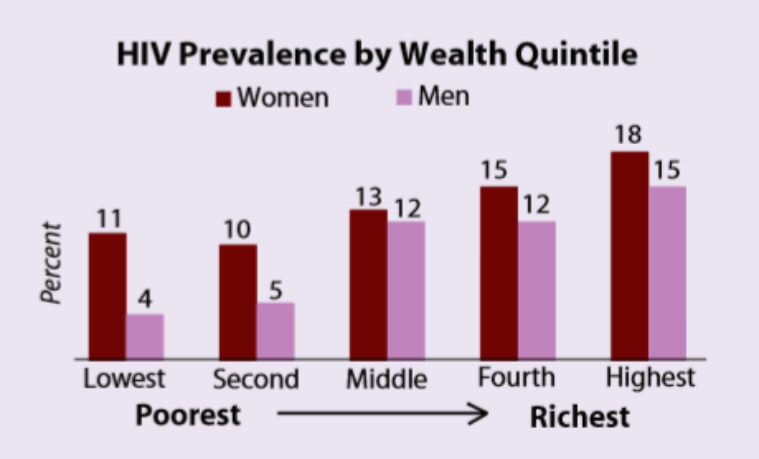
MLE

But, How Do You Develop a Very Strong Visualization from a Data Table?

Table 12.4 HIV prevalence by socioeconomic characteristics

Percentage HIV positive among women and men age 15-49 who were tested, by background characteristics, Malawi 2004

Background characteristic	Women		Men		Total	
	Percentage HIV positive	Number	Percentage HIV positive	Number	Percentage HIV positive	Number
Residence						
Urban	18.0	410	16.3	462	17.1	872
Rural	12.5	2,276	8.8	2,003	10.8	4,279
Region						
Northern	10.4	403	5.4	348	8.1	751
Central	6.6	1,032	6.4	994	6.5	2,026
Southern	19.8	1,251	15.1	1,122	17.6	2,373
District						
Blantyre	22.5	211	22.1	247	22.3	457
Kasungu	5.5	116	2.8	116	4.1	232
Machinga	14.9	99	8.2	86	11.8	185
Mangochi	21.4	136	19.9	108	20.8	244
Mzimba	6.4	178	3.9	157	5.2	336
Salima	9.5	74	8.2	59	8.9	133
Thyolo	23.1	145	18.6	126	21.0	271
Zomba	24.6	134	10.5	124	17.8	258
Lilongwe	a	352	a	398	a	750
Mulanje	23.3	117	15.1	86	19.8	202
Other districts	12.6	1,124	8.7	958	10.8	2,082
Education						
No education	13.6	667	9.2	295	12.3	962
Primary 1-4	12.3	690	6.5	550	9.7	1,240
Primary 5-8	13.2	958	10.8	957	12.0	1,916
Secondary+	15.1	370	12.9	662	13.7	1,032
Employment						
Currently working	14.6	1,545	13.4	1,360	14.0	2,904
Not currently working	11.6	1,141	6.3	1,105	9.0	2,246
Wealth quintile						
Lowest	10.9	455	4.4	300	8.3	755
Second	10.3	546	4.6	487	7.6	1,033
Middle	12.7	581	12.1	568	12.4	1,149
Fourth	14.6	595	11.7	564	13.2	1,159
Highest	18.0	508	14.9	546	16.4	1,054
Ethnicity						
Chewa	8.3	915	5.7	796	7.1	1,711
Tumbuka	8.8	280	5.1	231	7.1	511
Lomwe	20.0	450	17.9	433	19.0	884
Tonga	15.9	53	(9.2)	49	12.6	102
Yao	17.5	337	12.4	300	15.1	637
Sena	12.4	126	10.3	115	11.4	241
Nkonde	(19.1)	30	4.9	51	10.2	82
Ngoni	14.5	288	14.5	299	14.5	587
Other	17.3	204	9.2	190	13.4	394
Religion						
Catholic	13.8	634	10.5	552	12.3	1,185
CCAP	9.7	481	8.7	464	9.2	945
Anglican	17.7	56	(5.6)	47	12.1	103
Seventh Day Adventist/Baptist	12.1	158	16.5	175	14.4	334
Other Christian	13.8	1,025	10.0	890	12.0	1,915
Muslim	17.0	309	11.3	251	14.4	560
No religion	*	20	0.0	75	0.7	95
Total	13.3	2,686	10.2	2,465	11.8	5,150



Process:
Step 3

Developing a Data Visualization

Learning Activity 1 – Brief overview of scientific aspects of visual perception and an interactive activity (20 minutes)

Learning Activity 2 – Brief overview of the four key principles to follow when developing a data visualization:

- selection of the appropriate data graphic
- simplicity of style in presentation
- careful choice of pre-attentive attributes
- honesty

(40 minutes)

MLE

Process:
Step 3

Developing a Data Visualization (continued)

Learning Activity 3 – Brief review of all the topics presented yesterday (10 minutes)

Learning Activity 4 – Participants working in pairs/ groups will create a data visualization (by hand) that is intended for a designated, target audience, with justifications prepared of this visualization's content and design; participants will then exchange their visualization for that of another pair/group in order to evaluate it along these same considerations (80 minutes)

MLE

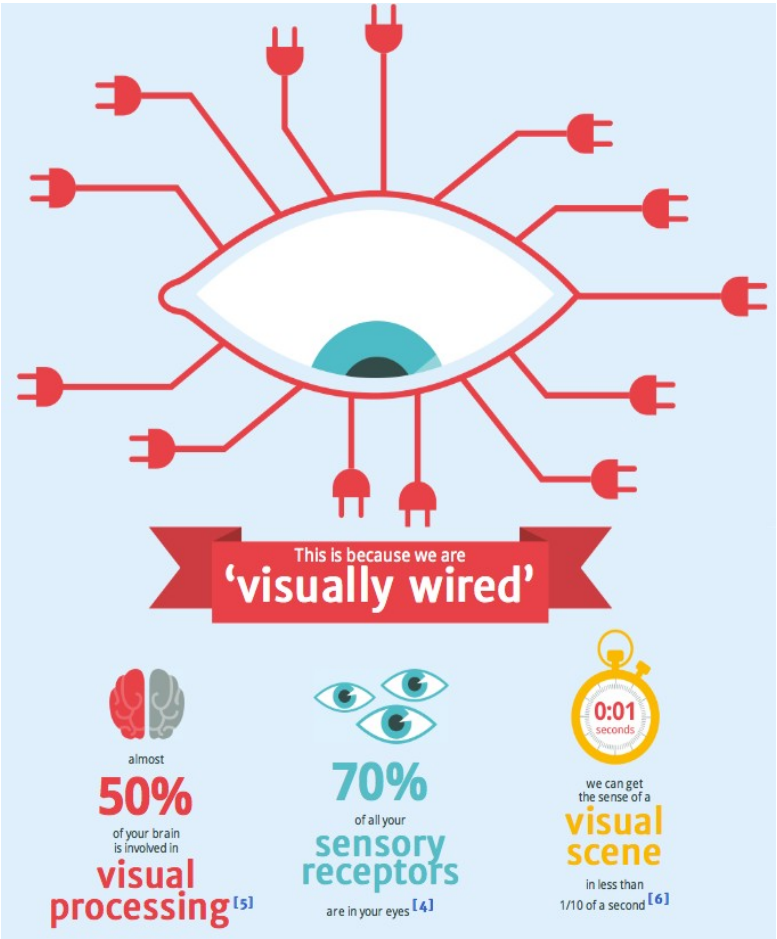
Developing a Data Visualization :
Aspects of Visual Perception

As objectives, at the end of this session, participants will be able to:

- Apply the principles of visual perception and the design principles for a data visualization when evaluating a data graphic (10 minutes)

MLE

Principles of Visual Perception and Human Cognition



Visual Perception (vision)

- Is processed at the visual cortex, a rear section of the brain section
- Works quickly and efficiently
- Provides an immediate visual sweeping of the world around us

Cognition (cerebral response to perceived images)

- Is processed mainly by the cerebral cortex, a front section of the brain section
- Works more slowly and less efficiently that visual perception

MLE

Significant Differences between Visual Perception and Cognition

- **Cognition**
  - Processing of **visual** information undertaken *consciously* by the brain
  - Brain function most often used when analyzing and reporting on data, since a focused attention is required
- **Visual Perception**
  - Processing of **visual** information undertaken *unconsciously* by the brain and one that is quicker than the processing of *verbal* information
  - Brain function underused when a targeted audience could be assisted in interpreting important information before cognition's slower, more analytical focus on the data

MLE

Lessons to Consider regarding Visual Perception

- In view of the vast quantity of information to share with others, data findings can be **conveyed more quickly and compellingly by a series of images rather than by text**

- Moreover, individuals can remember only 4 - 7 pieces of information at a time with their short-term memories, and the adopted use of principles related to visual perception could help **reinforce the key messages for a targeted audience**

- In sum, data visualization tips the balance towards **a greater exploitation of these principles of visual perception**

MLE

Visual Perception: Pre-Attentive Attributes

- Processing of visual information – undertaken unconsciously before the brain begins analytical processing of text and visual information– relies upon a processing of **pre-attentive** attributes of a visual image
- These pre-attentive attributes attract immediate attention when a visual image is perceived
  - **Color**
  - **Shape**
  - **Positioning**
  - **Movement**

Use these pre-attentive attributes in order to:

- facilitate visual sweeping of an image
- focus the attention of a targeted audience
- establish a visual hierarchy of elements for their display and comprehension
- assist in the remembrance of essential information

MLE

Visual Perception:
Pre-Attentive Attributes -- Color

How many number 5's are there below?

8	3	4	5	9	8	3	2	0	1
2	9	4	6	1	8	3	4	5	0
4	5	0	2	3	5	6	6	2	8
3	0	2	8	9	6	0	1	3	1
9	1	7	5	2	1	8	5	9	6

MLE

Visual Perception:
Pre-Attentive Attributes -- Color

8	3	4	**5**	9	8	3	2	0	1
2	9	4	6	1	8	3	4	**5**	0
4	**5**	0	2	3	**5**	6	6	2	8
3	0	2	**5**	9	6	0	1	3	1
9	1	7	8	2	1	8	**5**	9	6

MLE

Visual Perception:
Pre-Attentive Attributes -- Color

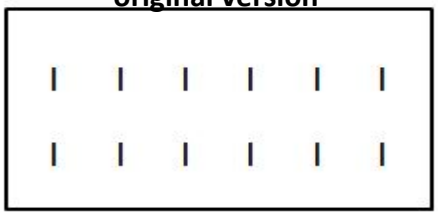
response: **6**

8	3	4	**5**	9	8	3	2	0	1
2	9	4	6	1	8	3	4	**5**	0
4	**5**	0	2	3	**5**	6	6	2	8
3	0	2	**5**	9	6	0	1	3	1
9	1	7	8	2	1	8	**5**	9	6

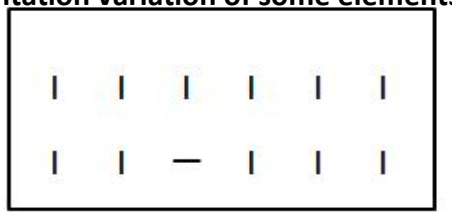
MLE

Visual Perception:
Pre-Attentive Attributes – Shape & Positioning (1)

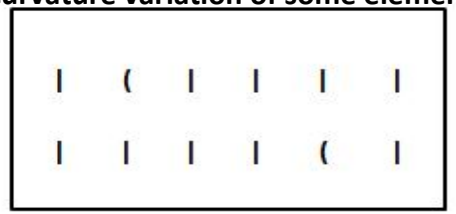
original version



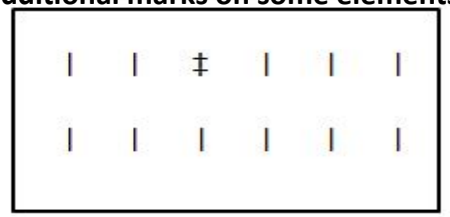
orientation variation of some elements



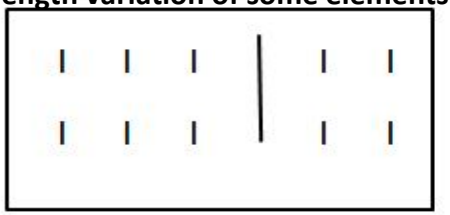
curvature variation of some elements



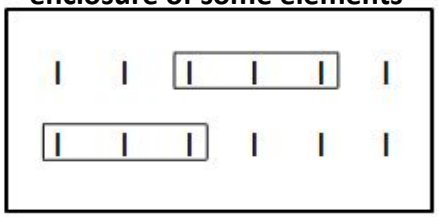
additional marks on some elements



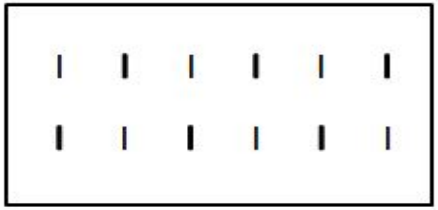
length variation of some elements



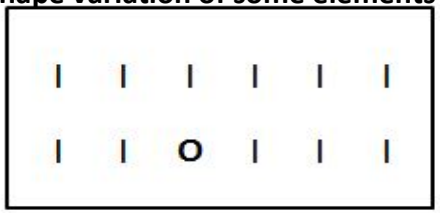
enclosure of some elements



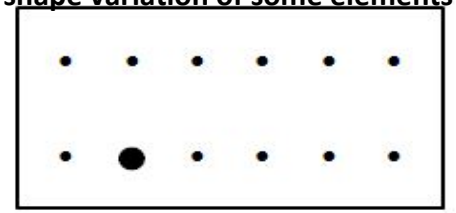
thickness variation of some elements



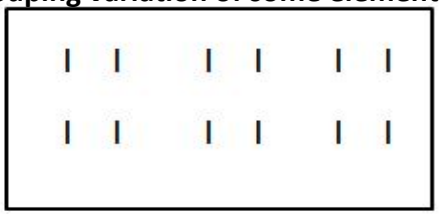
shape variation of some elements



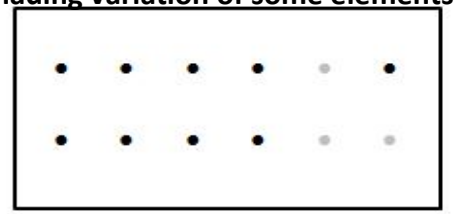
shape variation of some elements



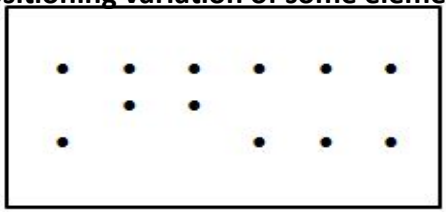
grouping variation of some elements



shading variation of some elements

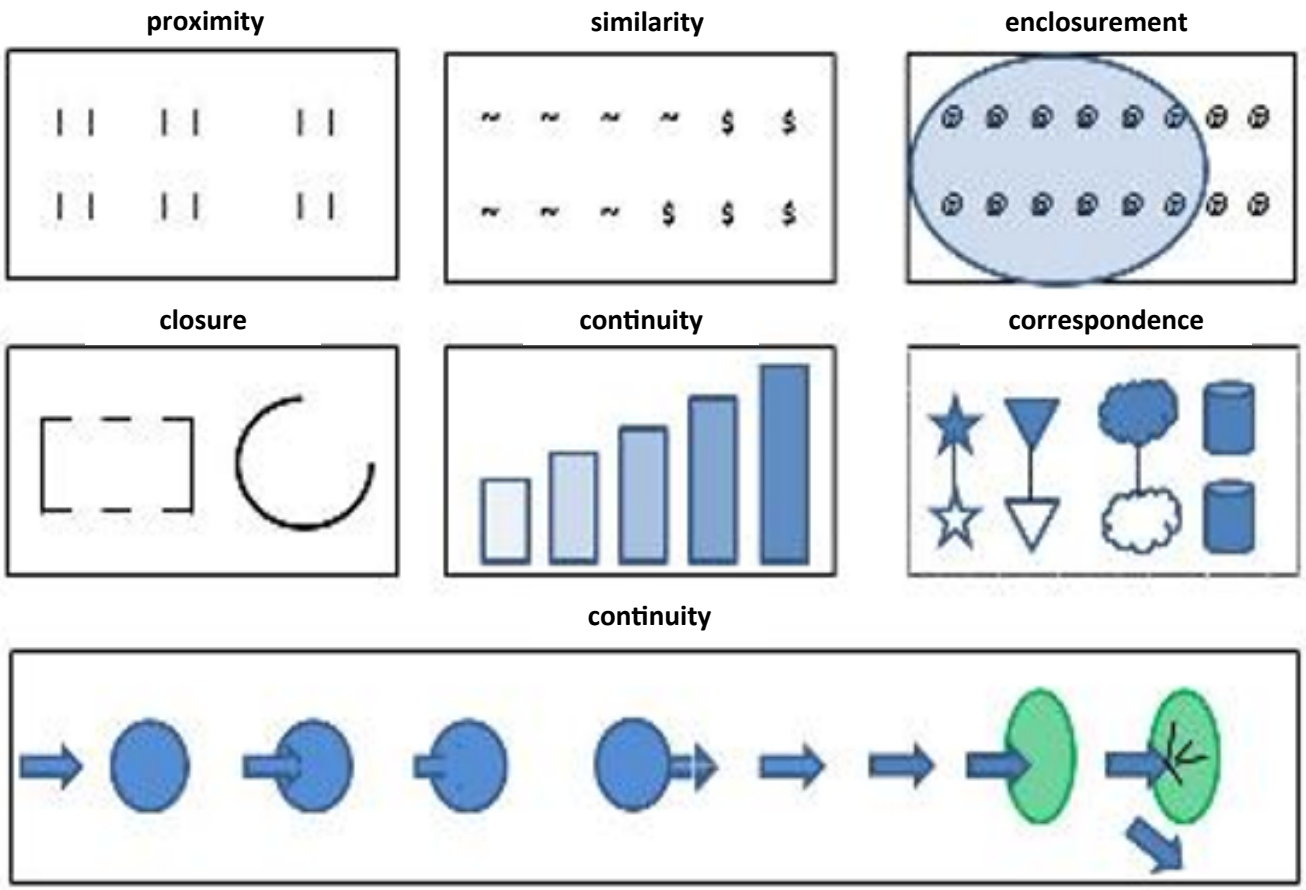


positioning variation of some elements



MLE

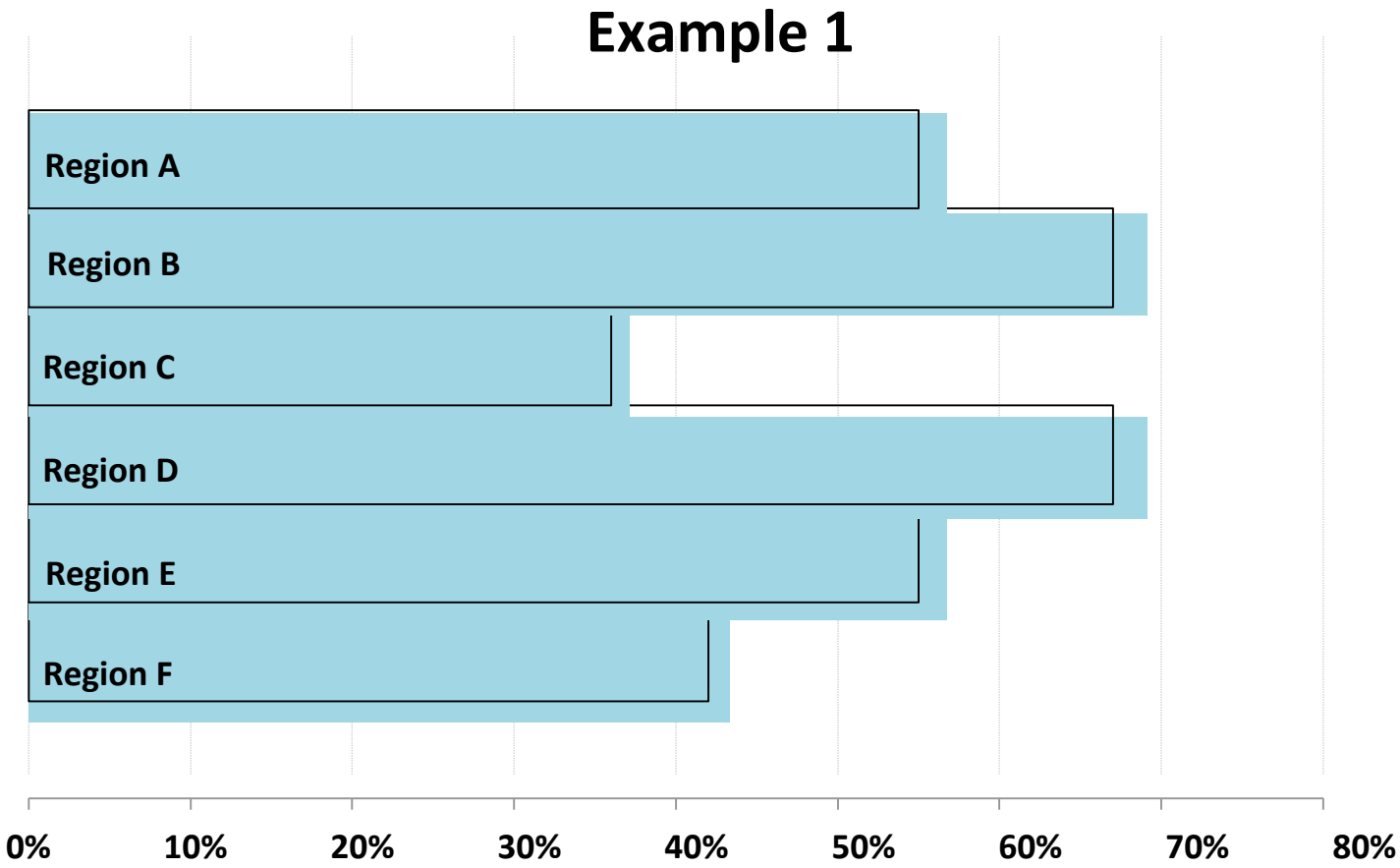
Visual Perception:
Pre-Attentive Attributes – Shape & Positioning (2)



Interactive Activity #1 (10 minutes)

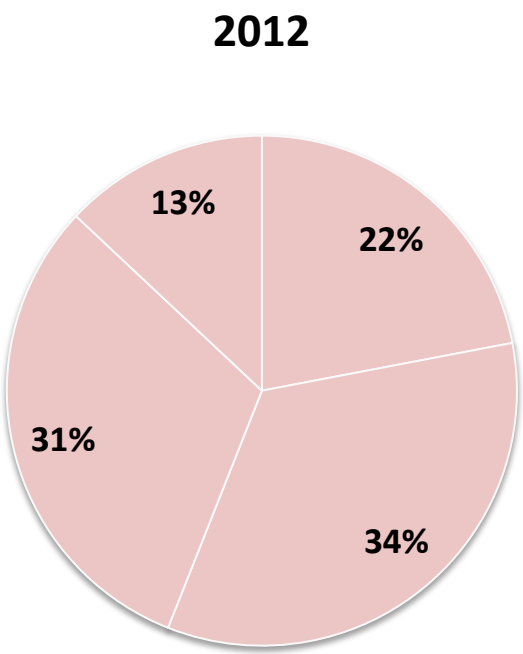
- Separate yourselves into pairs/groups in order to evaluate several examples of data graphics that **do not respect** the scientific aspects of visual perception
- Enumerate your reasons why these examples are « bad » and how they could be corrected

Interactive Activity #1 (continued)



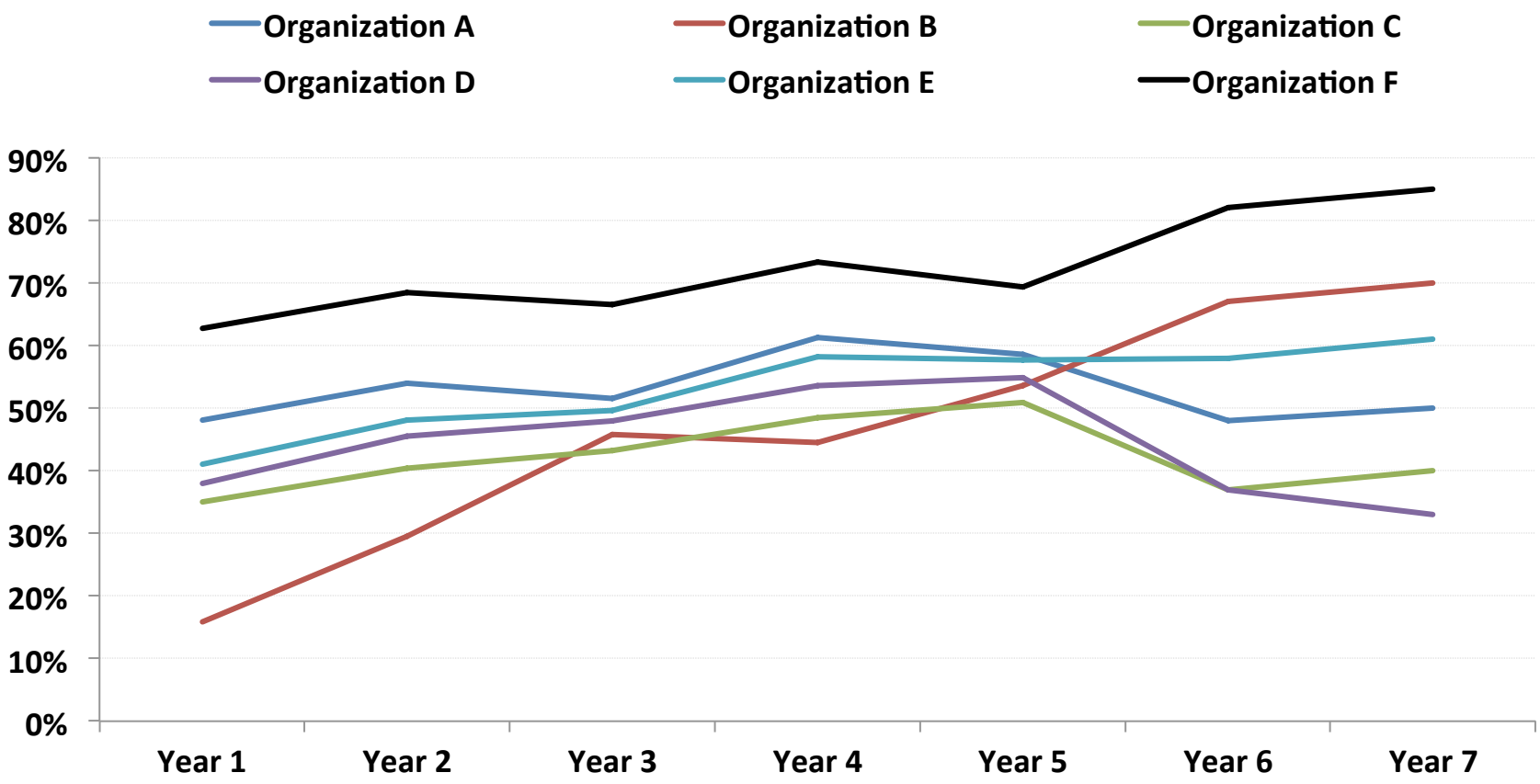
Interactive Activity #1 (continued)

Example 2



Interactive Activity #1 (continued)

Example 3



MLE

Developing a Data Visualization:

Aspects concerning the Four Principles pour the Development of a Data Visualization

As objectives, at the end of this session, participants will be able to

- Demonstrate an understanding of the four principles for developing a data visualisation

(40 minutes)

MLE

Four Principles related to the Development of a Data Visualization

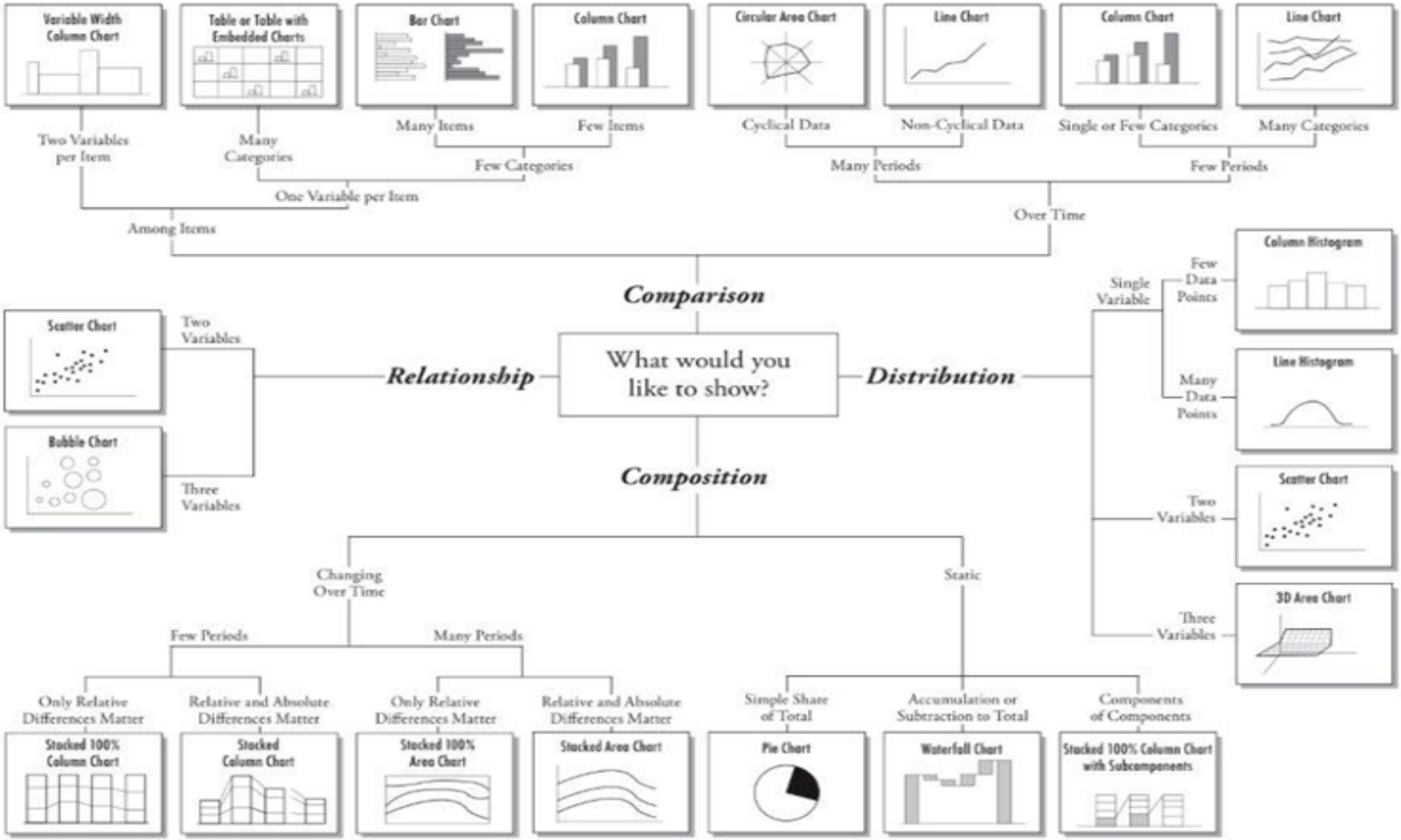
- Selection of the Appropriate Data Graphic
- Simplicity of Style in Presentation
- Careful Choice of Pre-Attentive Attributes
- Honesty

Four Principles related to the Development of a Data Visualization

- **Selection of the Appropriate Data Graphic**
- Simplicity of Style in Presentation
- Careful Choice of Pre-Attentive Attributes
- Honesty

Four Principles :

Selection of the Appropriate Data Graphic (1)



Four Principles :

Selection of the Appropriate Data Graphic (2)

- Ensure a good match between data findings and the data graphic to be used, since data can be shown as:
  - ***Distribution*** by percentage of variable values
  - ***Comparison*** among variables at one point in time or over the course of time
  - ***Relationship*** between variables
  - ***Composition*** of variables at one point in time or over the course of time

MLE

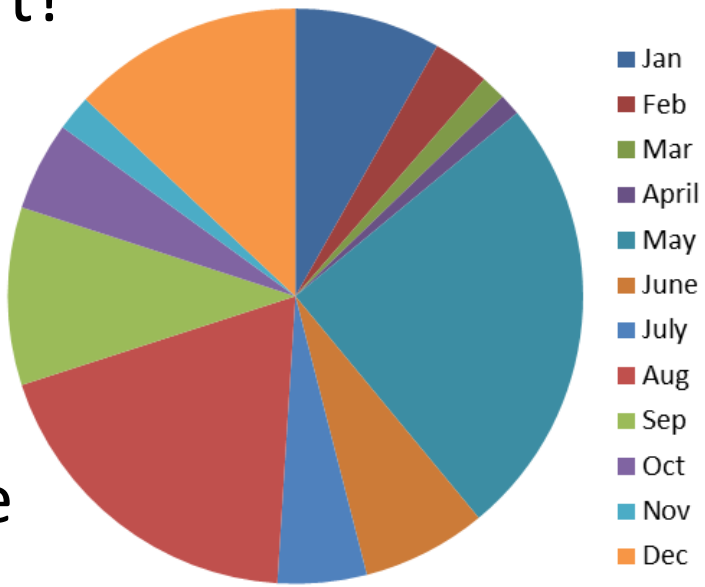
Four Principles :

Selection of the Appropriate Data Graphic (3)

What is the problem with a pie chart?

It is not easy to:

- Interpret the angles and the rounded edges
- Make comparisons among the wedges regarding their respective areas
- Make comparisons among too many wedges



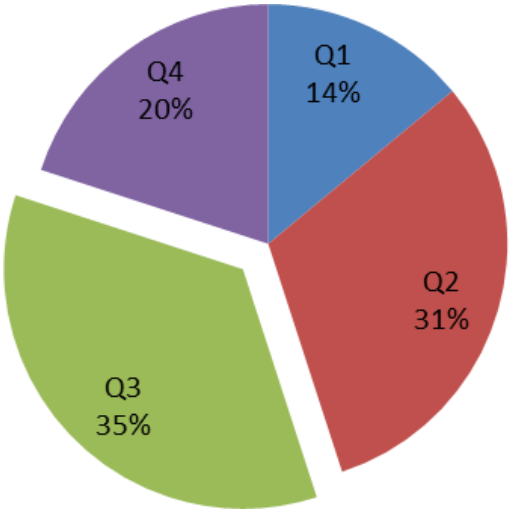
MLE

Four Principles :

Selection of the Appropriate Data Graphic (4)

But, the **use** of a pie chart **is justified** if these conditions are met:

- Number of wedges is limited to four to six
- Labeling is placed directly on each wedge
- Attention is focused on those wedges serving to transmit the message derived from the data; they should be separated from the others or distinguished in some other way
- Percentage of the wedges add up to 100%
- Pie chart is presented in two-dimensional form, *not three-dimensional*



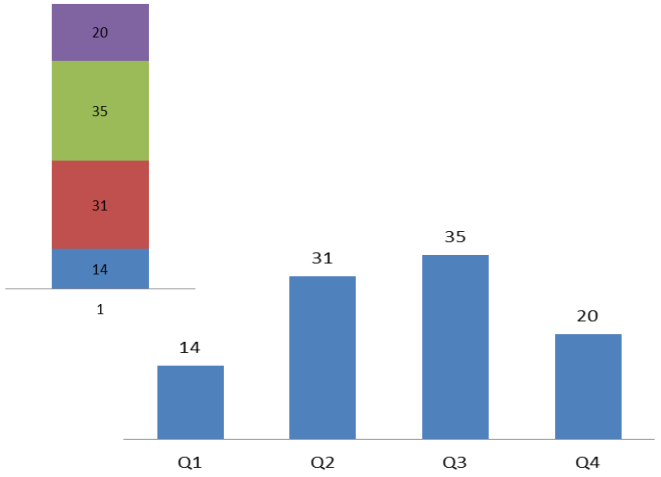
MLE

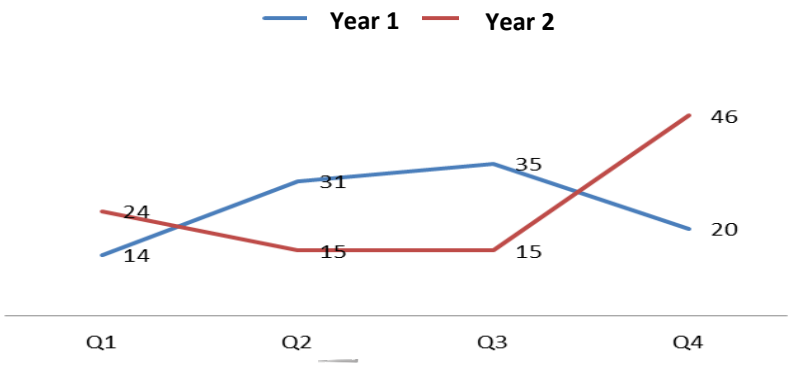
Four Principles :

Selection of the Appropriate Data Graphic (5)

On the other hand, there are other data graphics that could serve as an alternative to a pie chart:

- To show a distribution by percentage of a whole
- To show a comparison among categories
- To show variation over the course of time





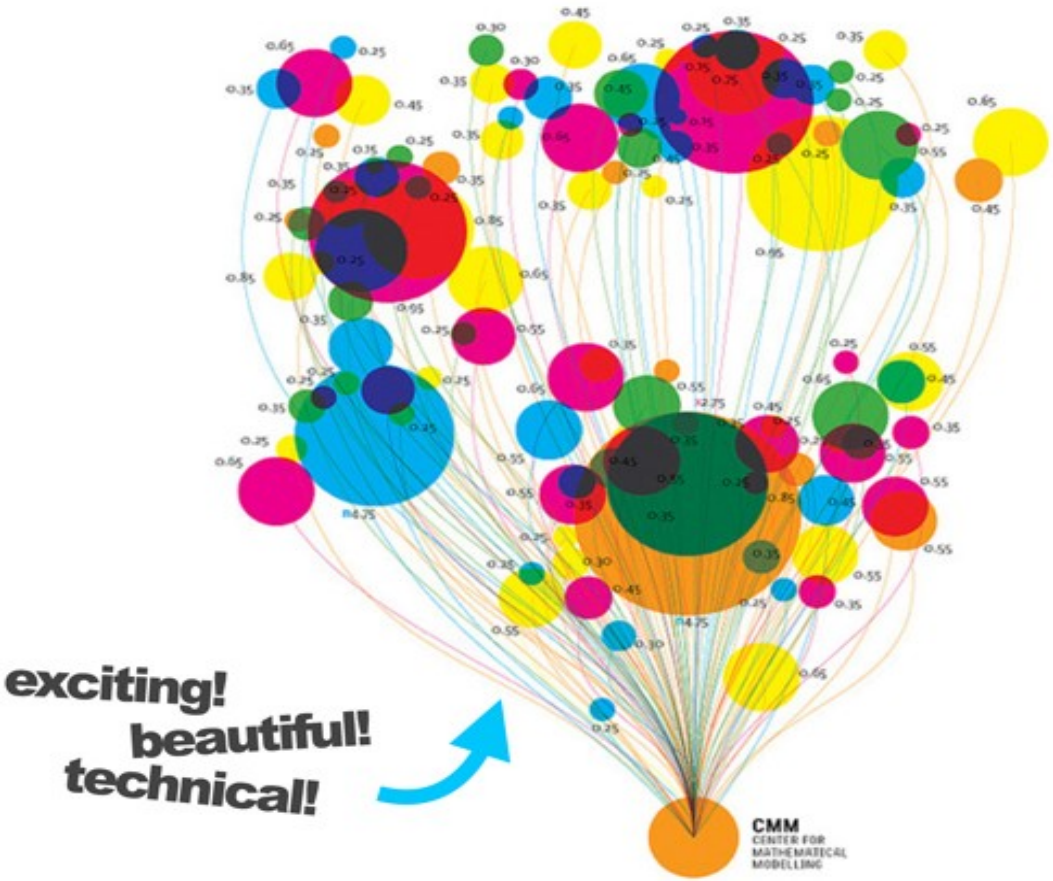
MLE

Four Principles related to the Development of a Data Visualization

- Selection of the Appropriate Data Graphic
- **Simplicity of Style in Presentation**
- Careful Choice of Pre-Attentive Attributes
- Honesty

MLE

But, What Does it Mean?

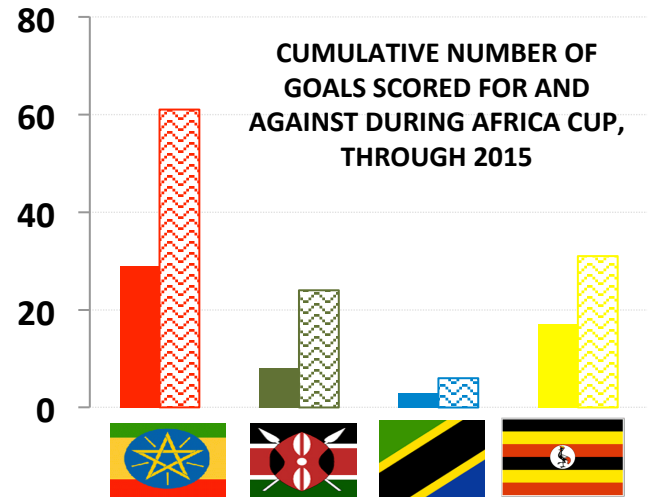


MLE

Four Principles :

Simplification of a Data Graphic (1)

- Data graphics should do without ornamentation
  - Avoid using clipart and other images for merely prettying up a data graphic
  - Nonetheless, take advantage of certain other sorts of images, like icons, in order to reenforce the interpretation of data in a graphic and to reduce the need for text labels



MLE

Four Principles :

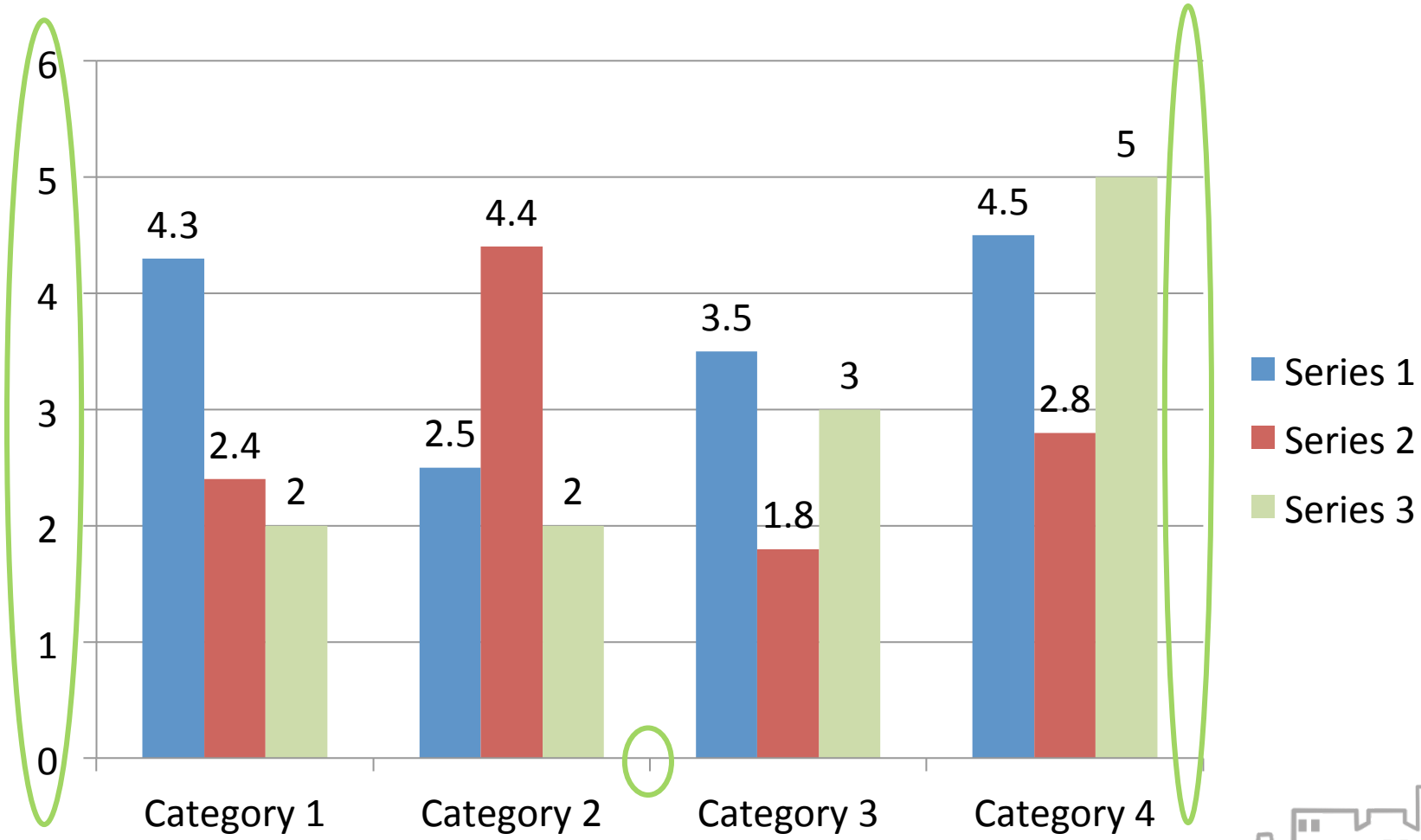
Simplification of a Data Graphic (2)

- Remove tick marks from the axes, unless they are truly necessary
- Change grid lines to a muted color, if they are to remain in a data graphic
  - Use light gray — not black — as the color
- Remove all borders around a data graphic
- Use only a single horizontal and vertical axis
- Limit the number of words in a title to six to twelve

MLE

Four Principles :

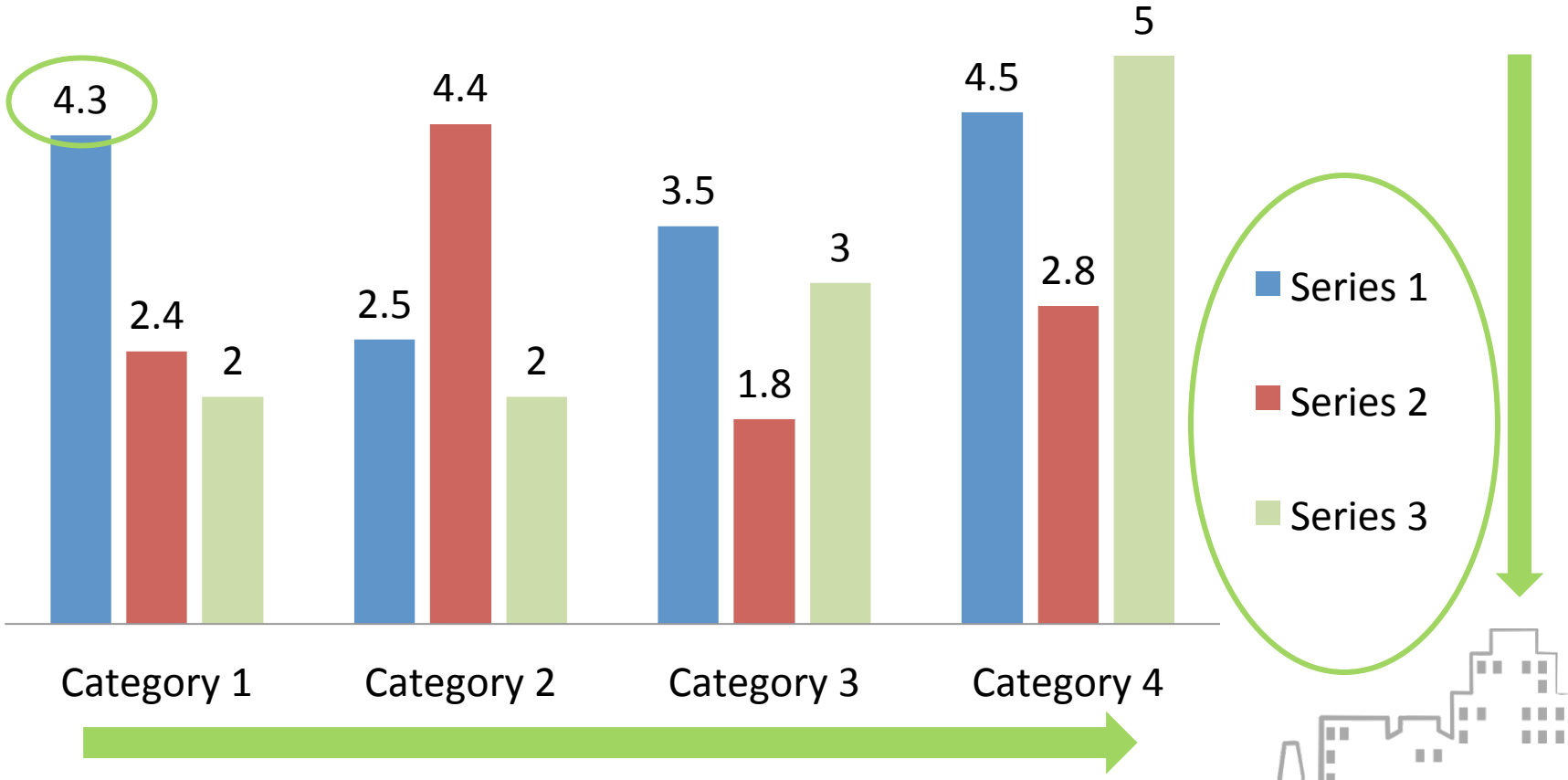
Simplification of a Data Graphic – First Example (1)



MLE

Four Principles :

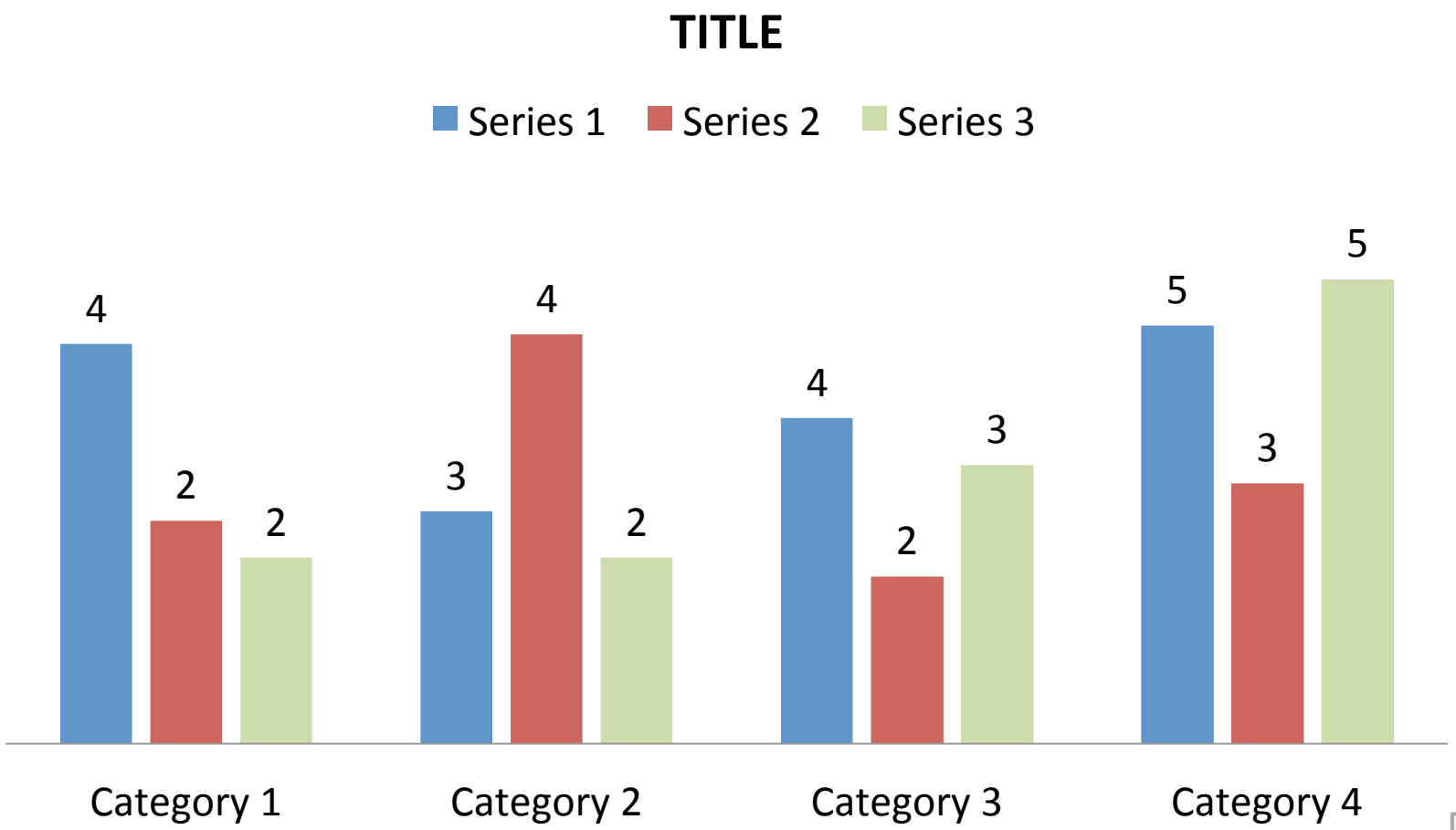
Simplification of a Data Graphic – First Example (2)



MLE

Four Principles :

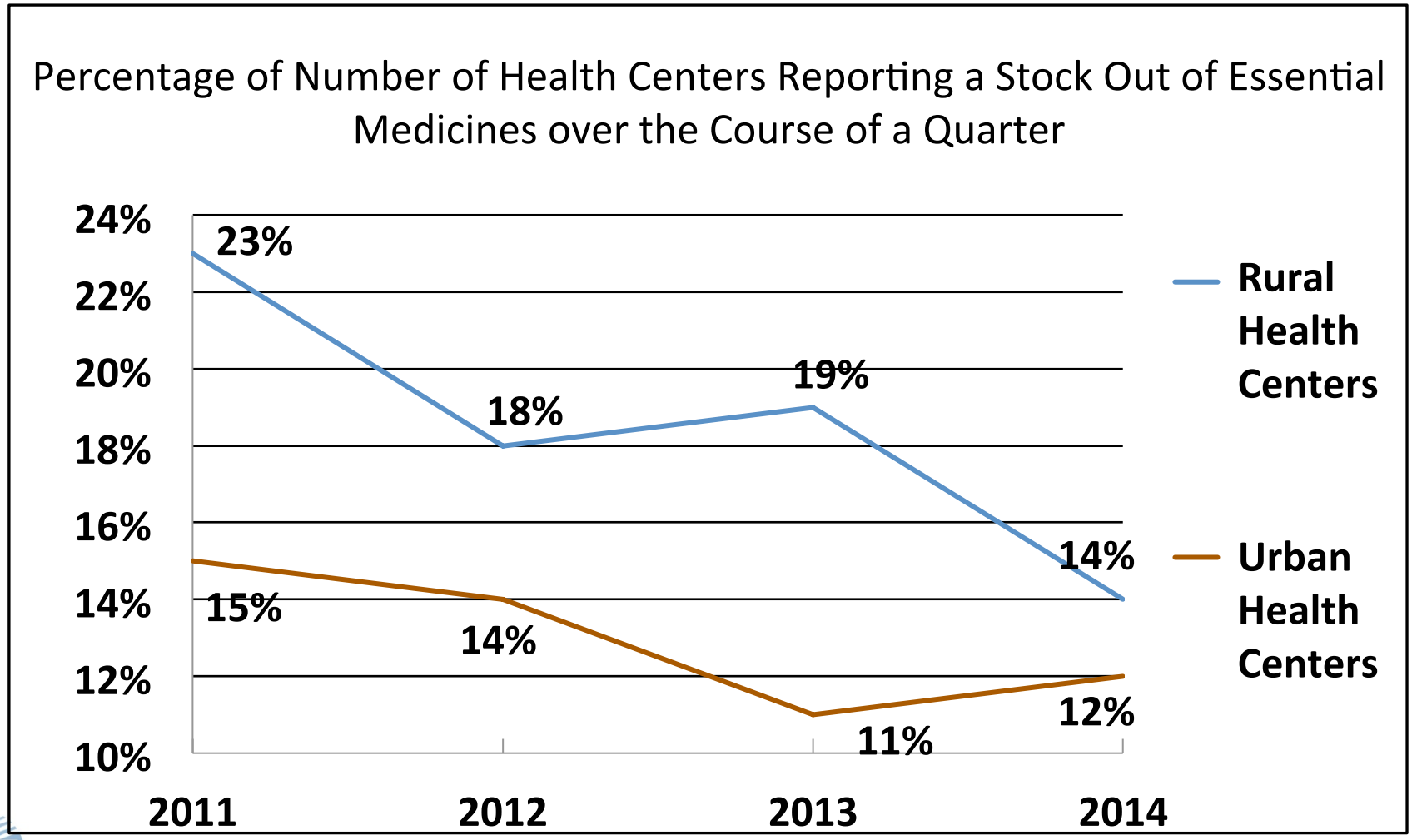
Simplification of a Data Graphic – First Example (3)



MLE

Four Principles :

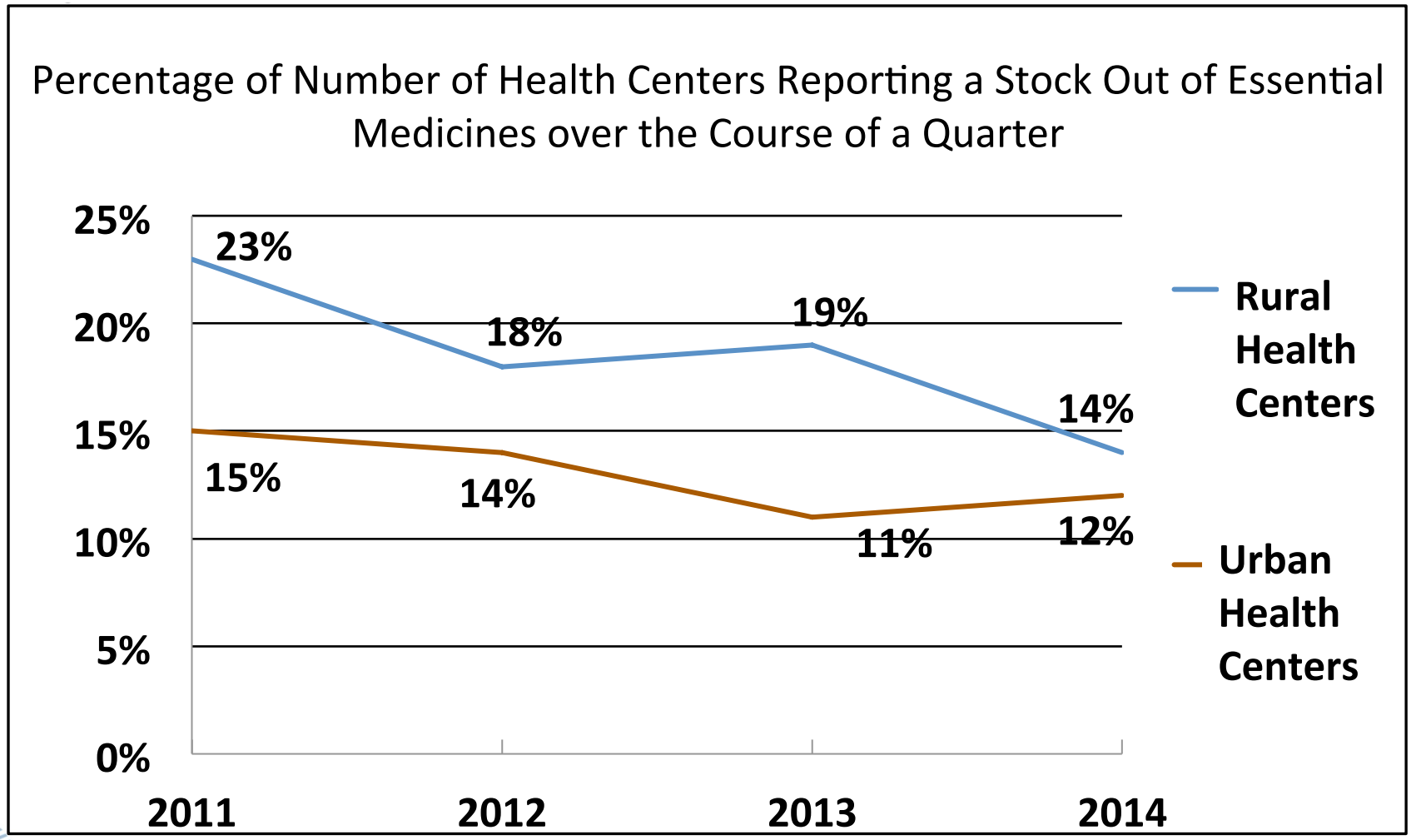
Simplification of a Data Graphic – Second Example (1)



MLE

Four Principles :

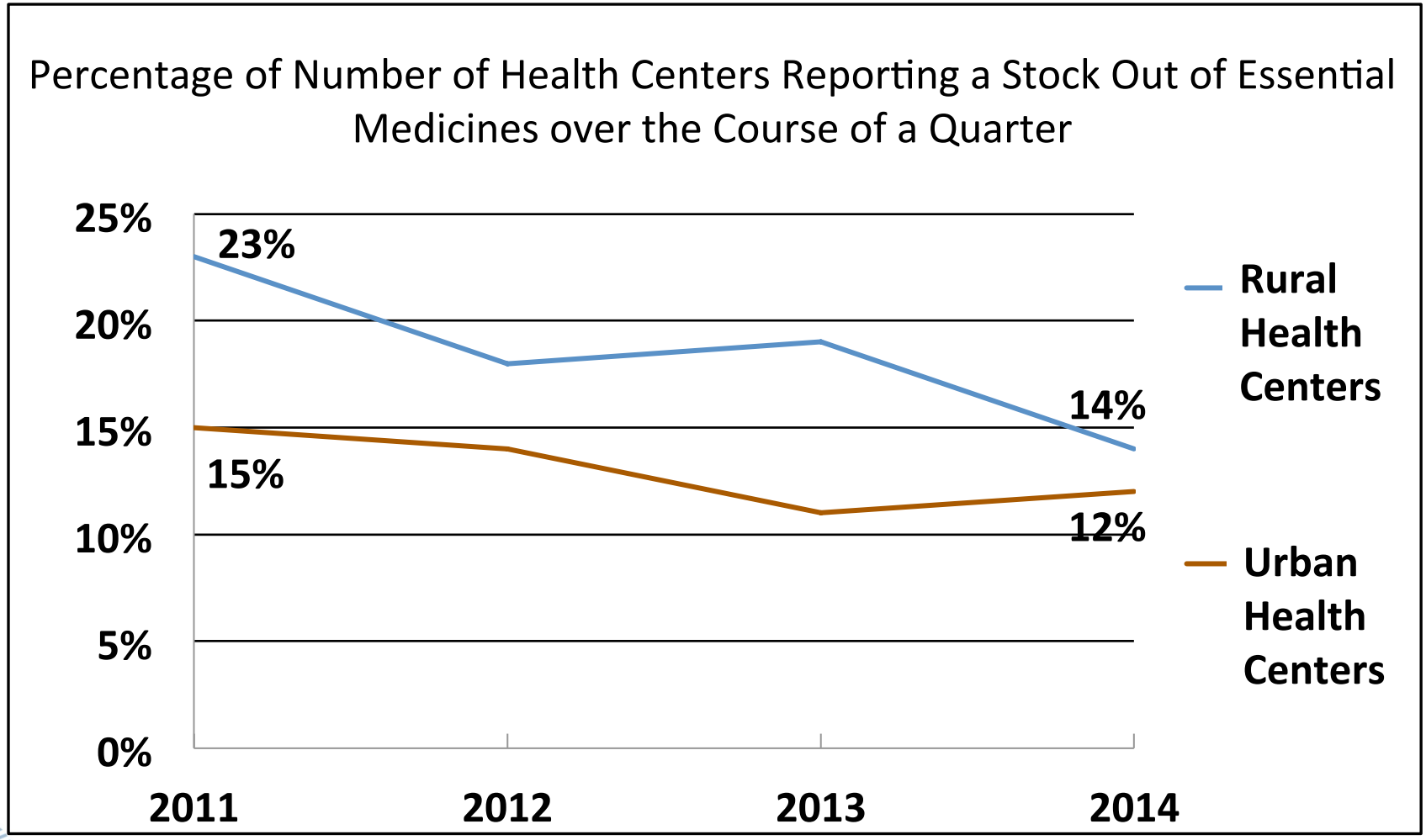
Simplification of a Data Graphic – Second Example (2)



MLE

Four Principles :

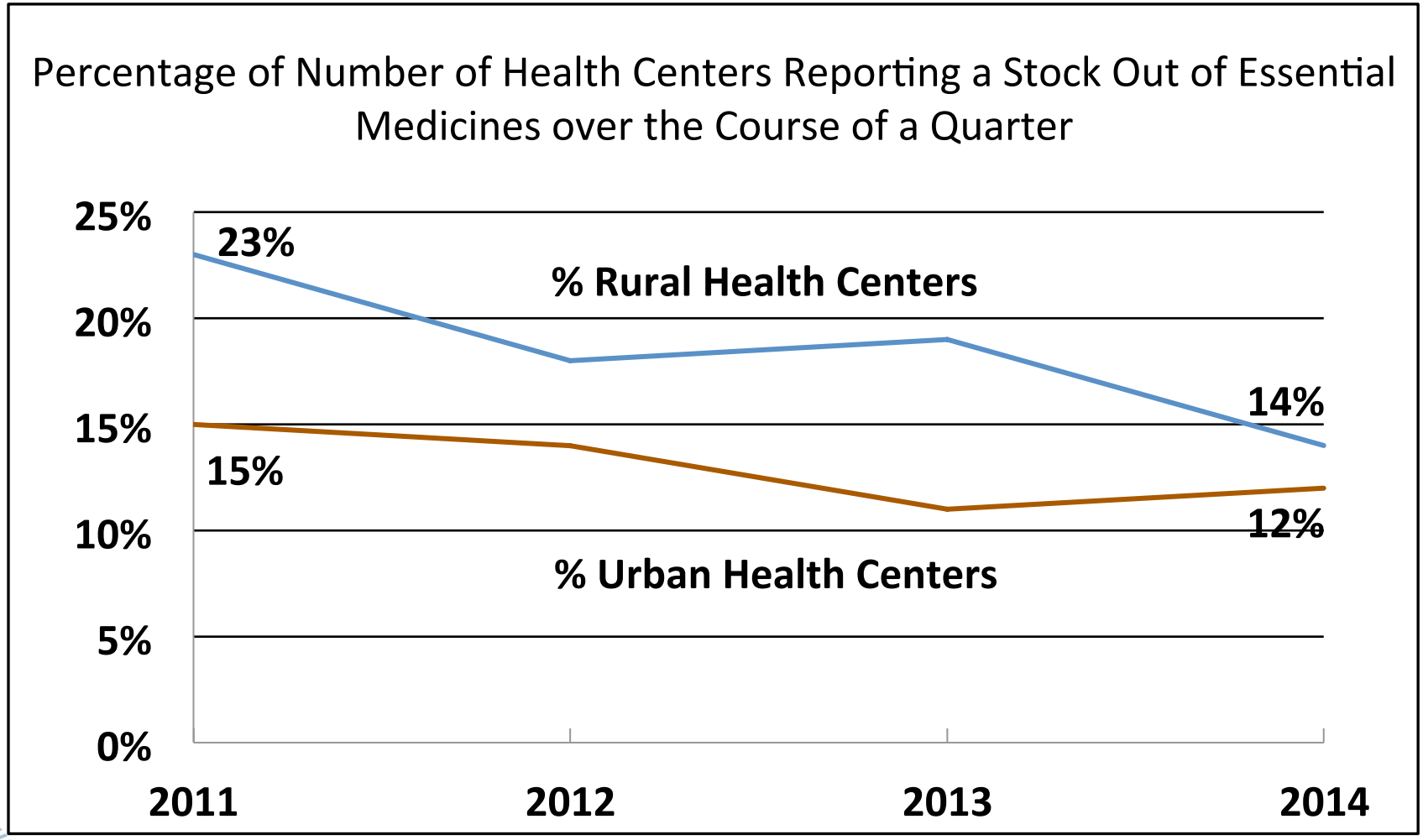
Simplification of a Data Graphic – Second Example (3)



MLE

Four Principles :

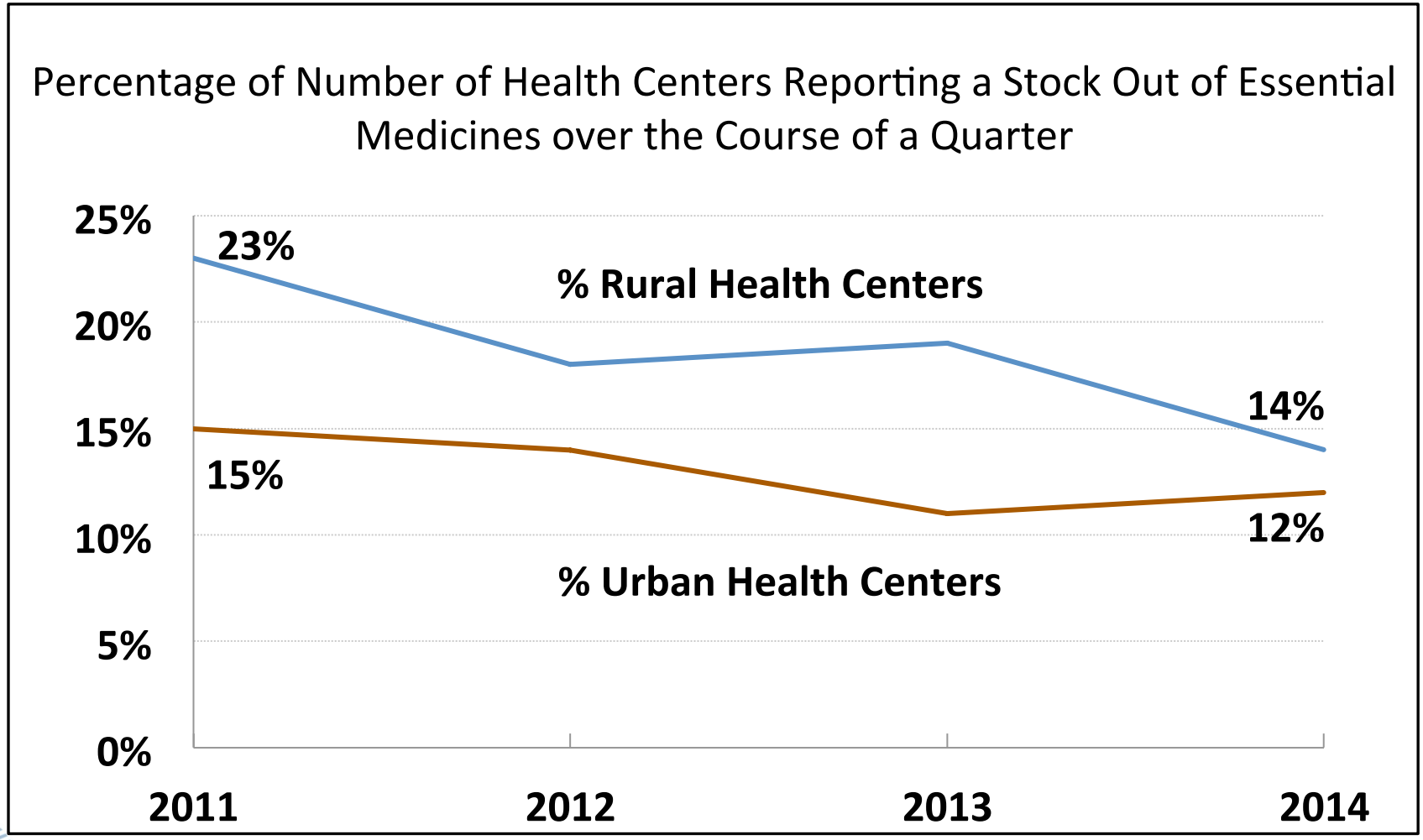
Simplification of a Data Graphic – Second Example (4)



MLE

Four Principles :

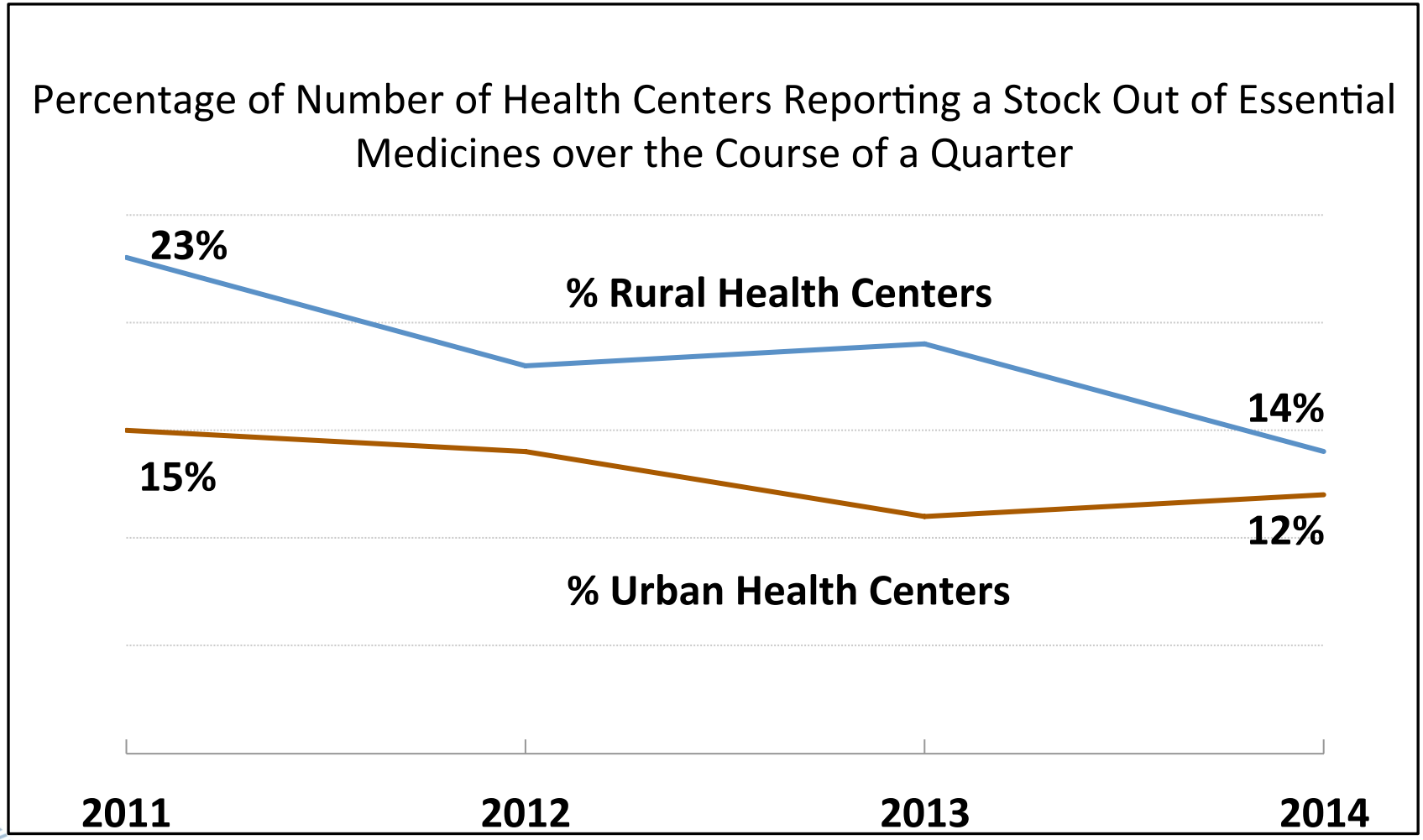
Simplification of a Data Graphic – Second Example (5)



MLE

Four Principles :

Simplification of a Data Graphic – Second Example (6)

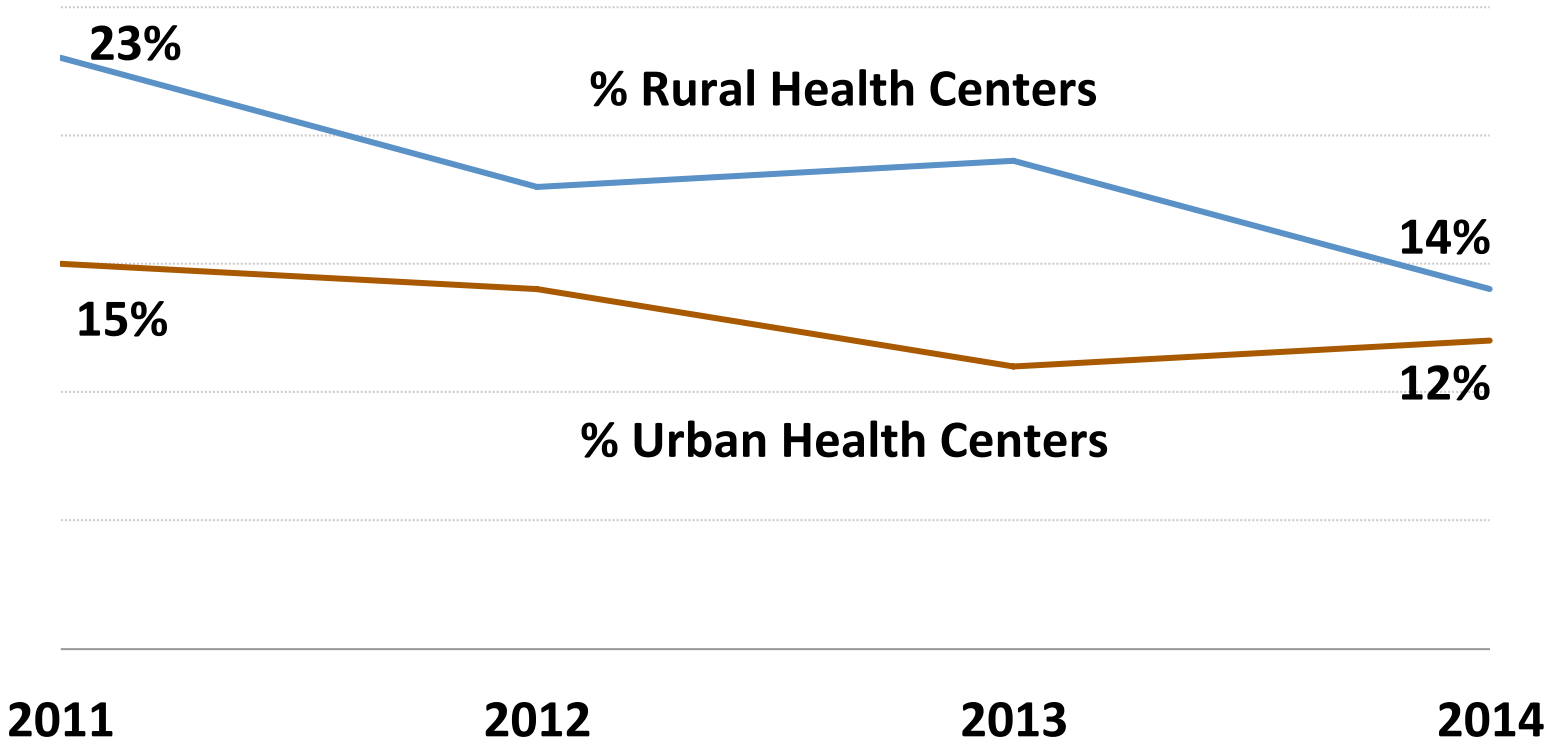


MLE

Four Principles :

Simplification of a Data Graphic – Second Example (7)

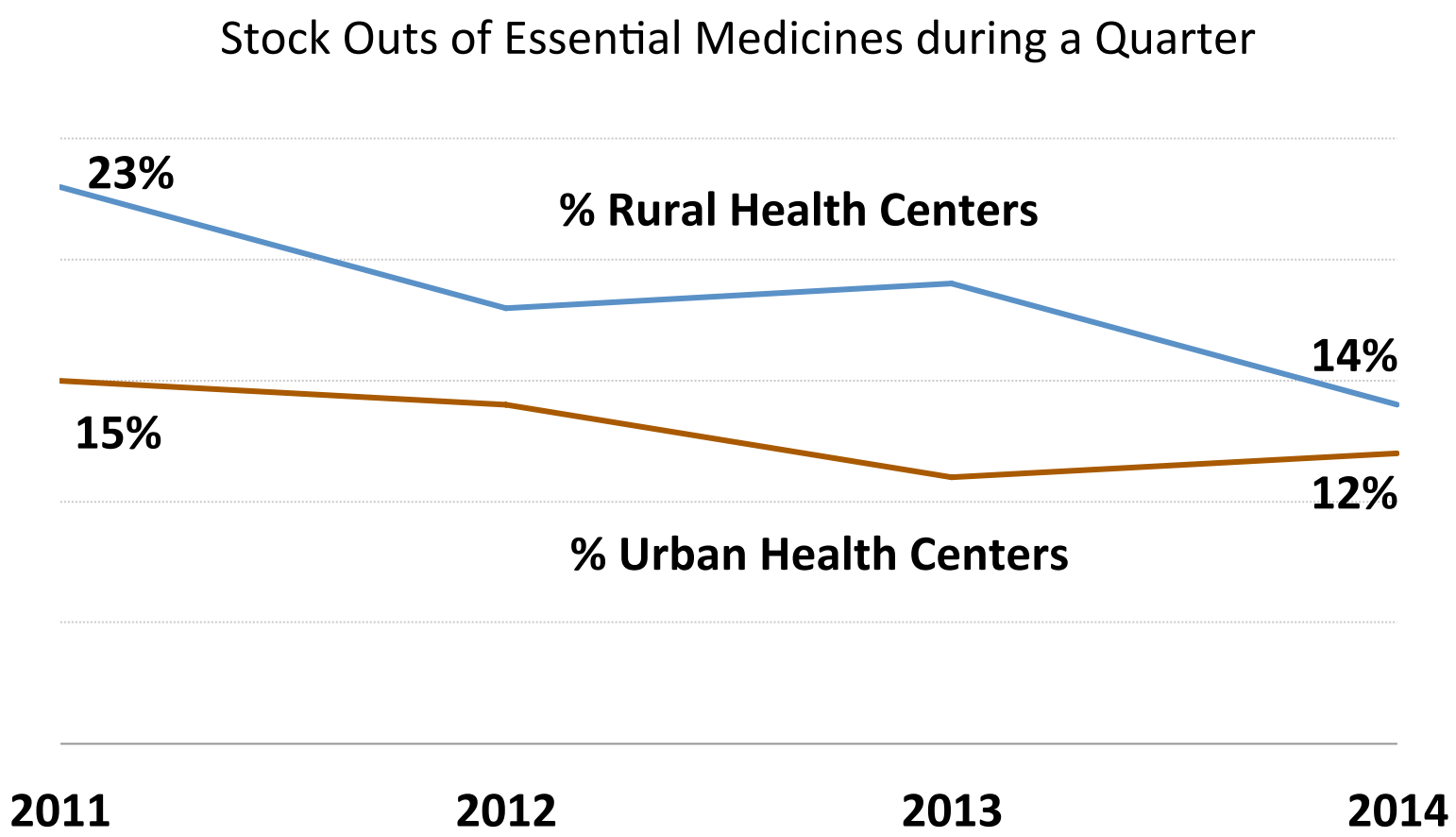
Percentage of Number of Health Centers Reporting a Stock Out of Essential Medicines over the Course of a Quarter



MLE

Four Principles :

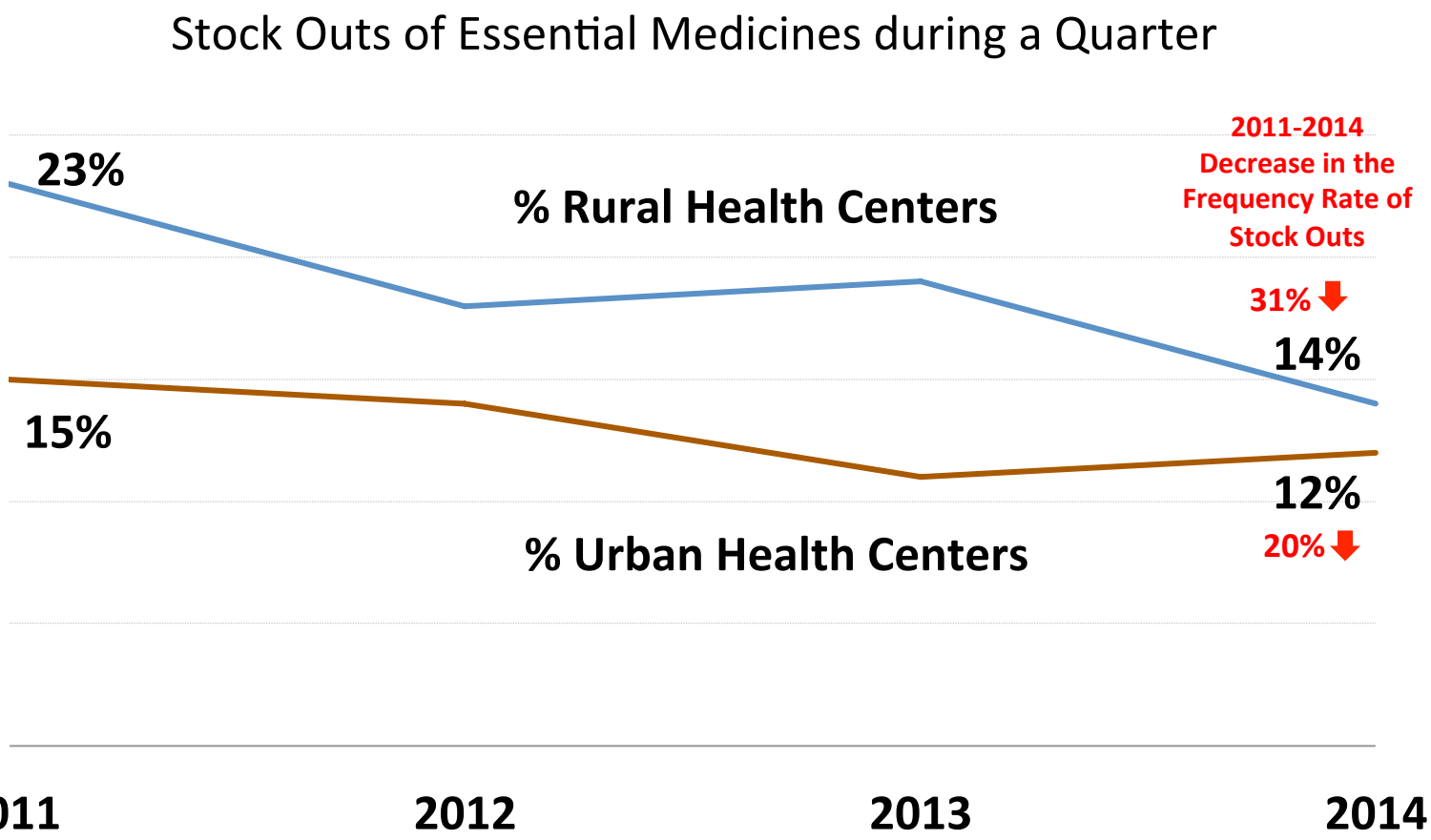
Simplification of a Data Graphic – Second Example (8)



MLE

Four Principles :

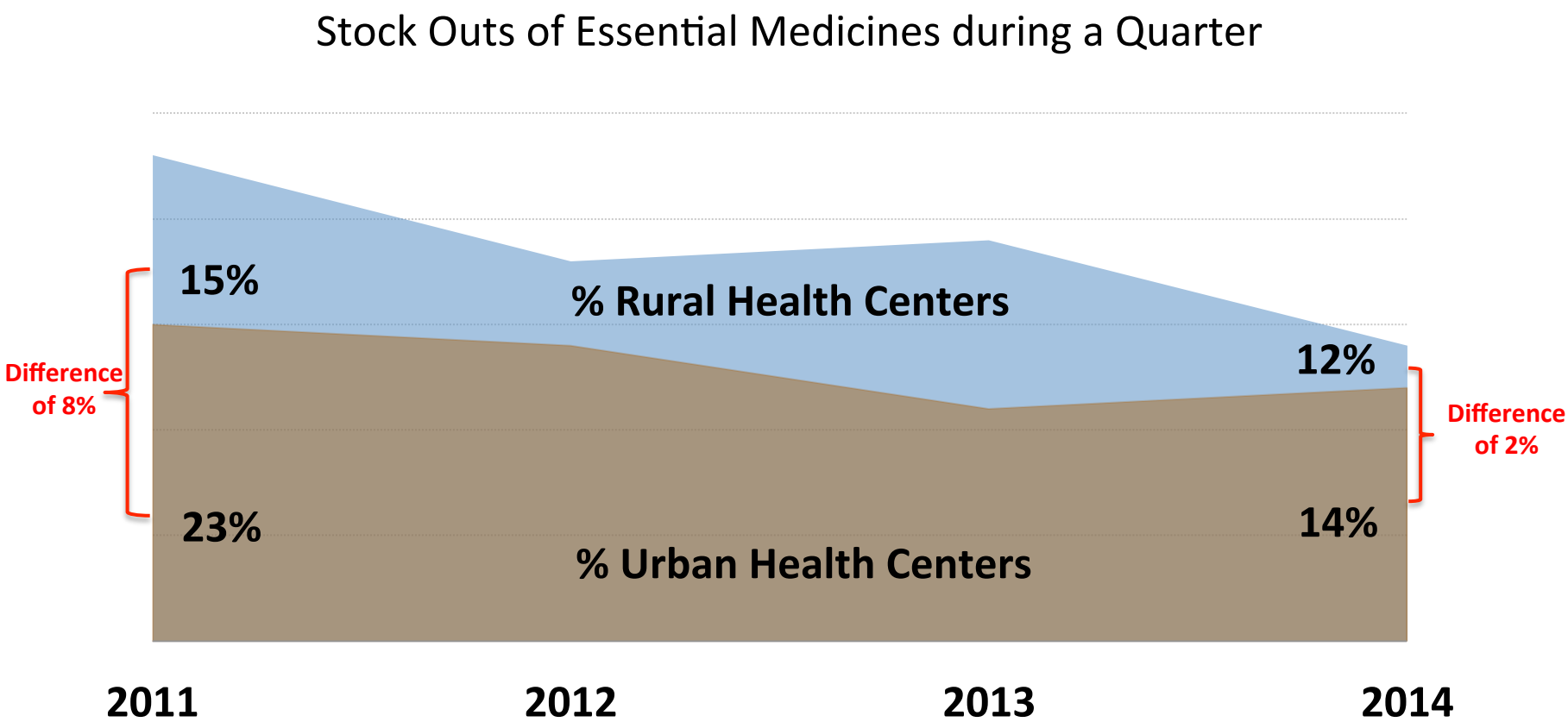
Simplification of a Data Graphic – Second Example (9)



MLE

Four Principles :

Simplification of a Data Graphic – Second Example (10)



MLE

Four Principles related to the Development of a Data Visualization

- Selection of the Appropriate Data Graphic
- Simplicity of Style in Presentation
- **Careful Choise of Pre-Attentive Attributes**
- Honesty

MLE

Four Principles :

Careful Choice of Pre-Attentive Attributes

- Attract immediate attention when a visual image is perceived by means of these pre-attentive attributes
  - **Color**
  - **Shape**
  - **Positioning**
  - **Movement**

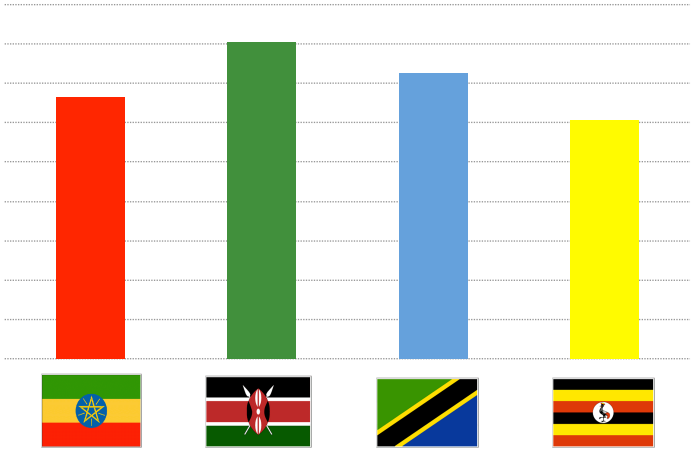
- Ensure that the design and elements of a data visualization serve well in eliciting a clear and rapid understanding of data findings

MLE

Four Principles:

Careful Choice of Pre-Attentive Attributes – **Color** (1)

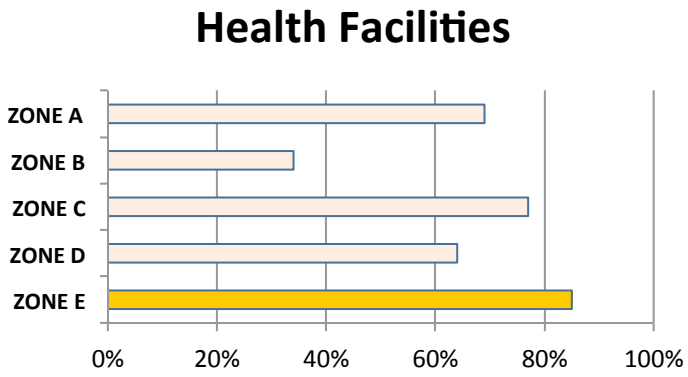
- Color schemes come from design considerations or the colors are intentionally chosen for specific reasons:
  - Facilitates the highlighting of trends presented by the data
  - Elicits an association to a branded image that is related to the study subject represented by the data



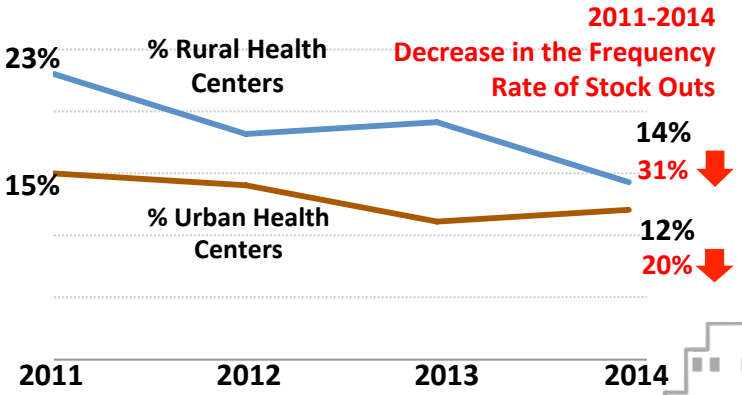
Four Principles:

Careful Choice of Pre-Attentive Attributes – **Color** (2)

- Use a different and catchy color in order to focus attention on the findings of meaningful data; use a muted color for the other data elements



- Similarly, use a striking color to draw immediate attention to text that provides explanation of meaningful findings from the data

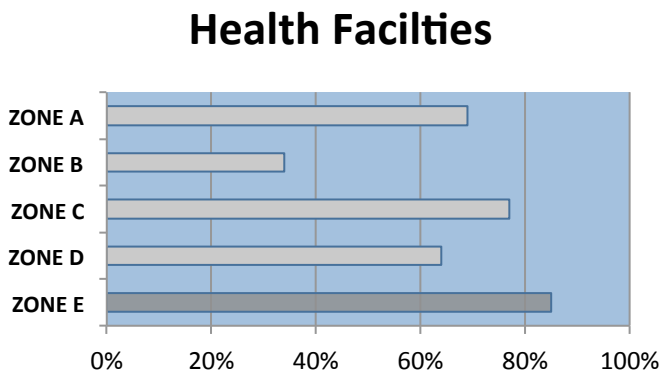


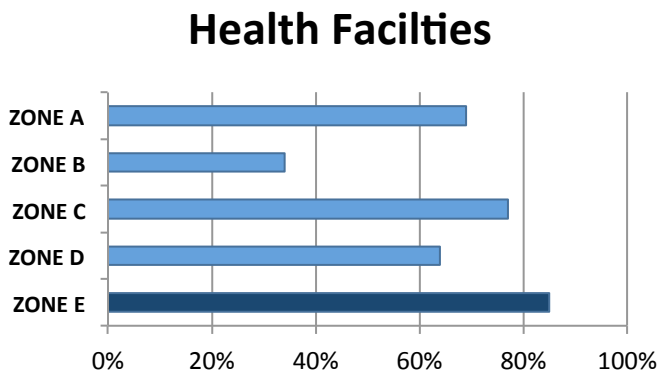
MLE

Four Principles:

Careful Choice of Pre-Attentive Attributes – **Color** (3)

- Colors of the text font and data elements should be in sufficient contrast to the background color of a data graphic
  - Interpretation of a data visualisation is greatly facilitated when text color is black or in a dark color while the background color is white or transparent

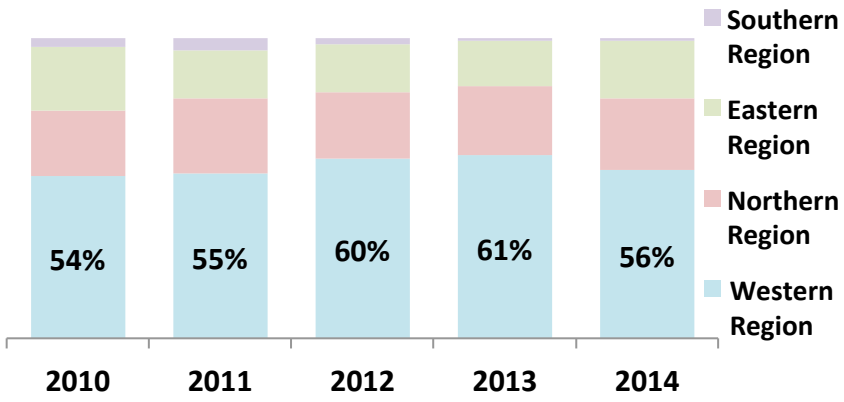




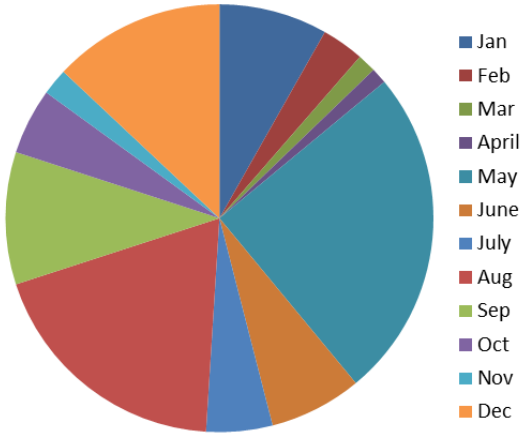
Four Principles:

Careful Choice of Pre-Attentive Attributes – **Color** (4)

- Use consistently a single color for the graphical display of a data variable category



- Avoid using more than six colors in a data graphic, since the appearance of multiple colors can distract the viewers

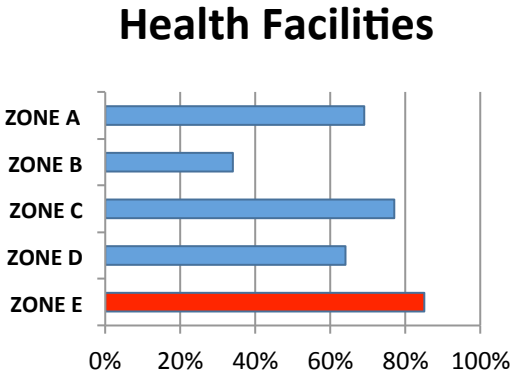


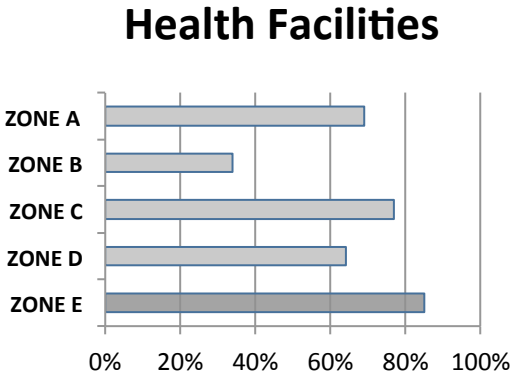
MLE

Four Principles:

Careful Choice of Pre-Attentive Attributes – **Color** (5)

- Avoid the risk of finding a lessened effectiveness with data visualizations designed in color:
  - Check for differences between colors after photocopying in black and white to ensure that there still remains differences in tone with which the trends shown by the data had been highlighted
  - Foresee the possibility that color copies of data visualizations could be reproduced by others in black and white



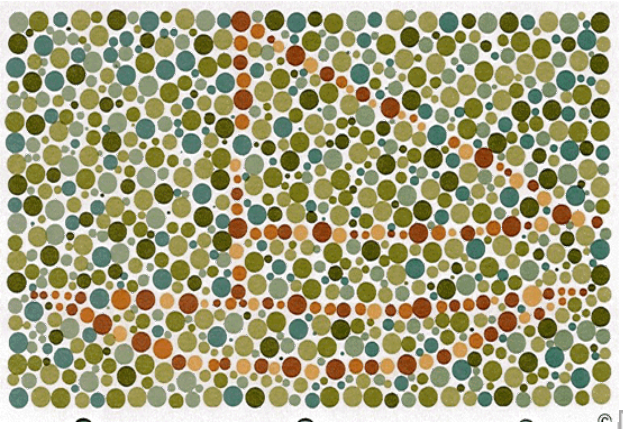


MLE

Four Principles:

Careful Choice of Pre-Attentive Attributes – **Color** (6)

- Avoid the use of color combinations listed below that may pose difficulties in their differentiation to those who are colorblind and those who are often distracted by the contrast of these colors:
  - **green** and **red**
  - **green** and **blue**
  - **green** and **gray**
  - **green** and **brown**
  - **green** and **black**
  - **light green** and **yellow**
  - **blue** and **yellow**
  - **blue** and **purple**
  - **blue** and **gray**

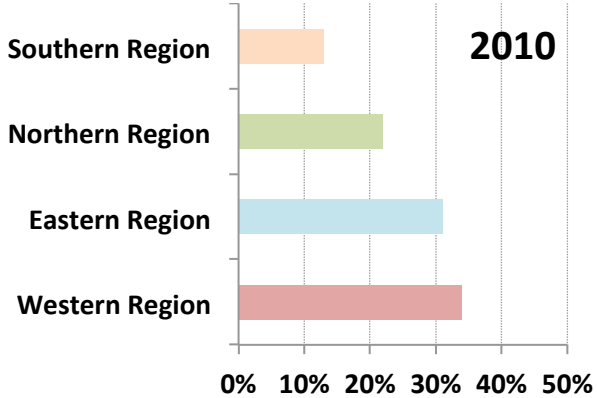


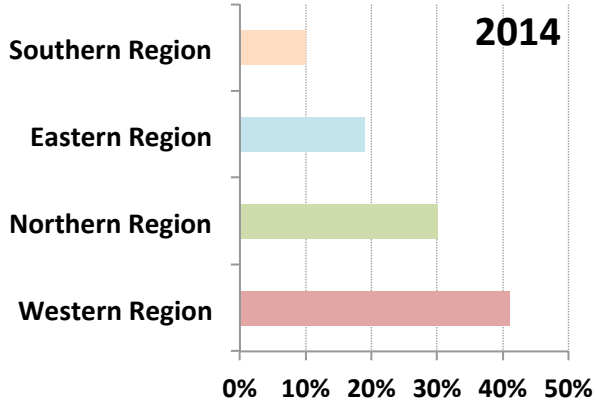
MLE

Four Principles:

Careful Choice of Pre-Attentive Attributes – **Shape** (1)

- Use a easily understandable method for differentiating data sets within a data graphic
  - Data should be sorted by an ordering system that is reasonable and logical, for example:
    - decreasing rate of frequency
    - increasing size of magnitude
    - alphabetic order
    - chronological order





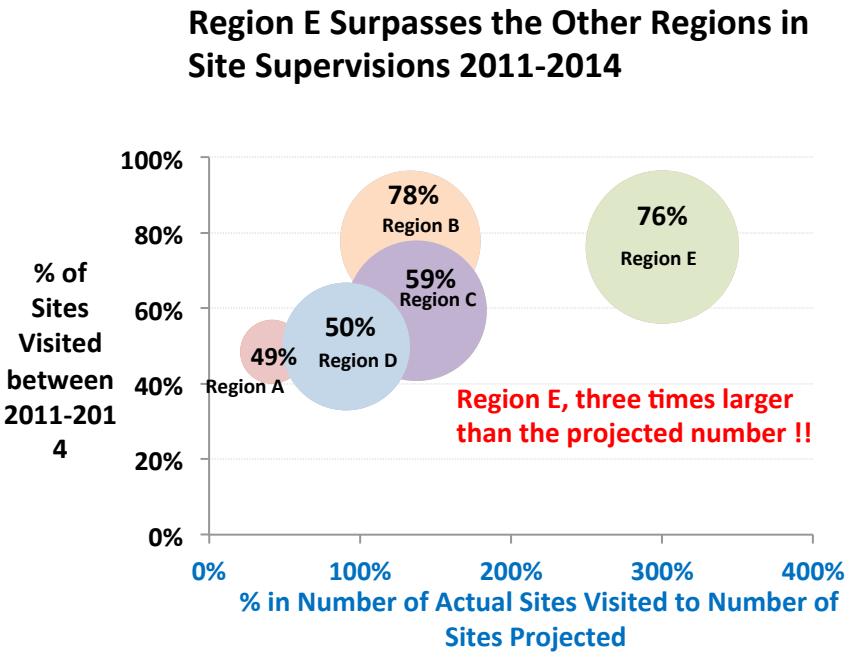
MLE

Four Principles:

Careful Choice of Pre-Attentive Attributes – **Shape** (2)

- Use the title of a data graphic or text boxes in order to highlight for the readers the most significant findings from the data

- Make sure that the font size of the text is legible and hierarchical (larger size indicating greater meaningfulness within the data graphic)



Title

Sub-Title

Comments

Text Label

Axis Label

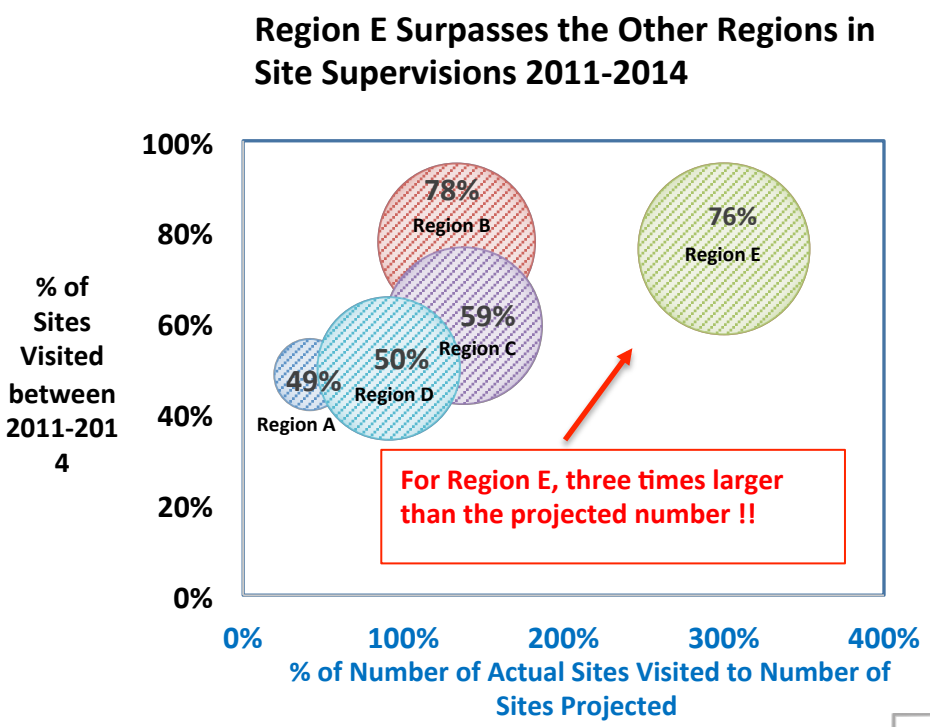
Footnotes

MLE

Four Principles:

Careful Choice of Pre-Attentive Attributes – **Shape** (3)

- Avoid using graphical elements that are unnecessary for a good interpretation of the data and that only serve to prettify or decorate the data graphic, which is often the case with arrows, boxes, and shadings

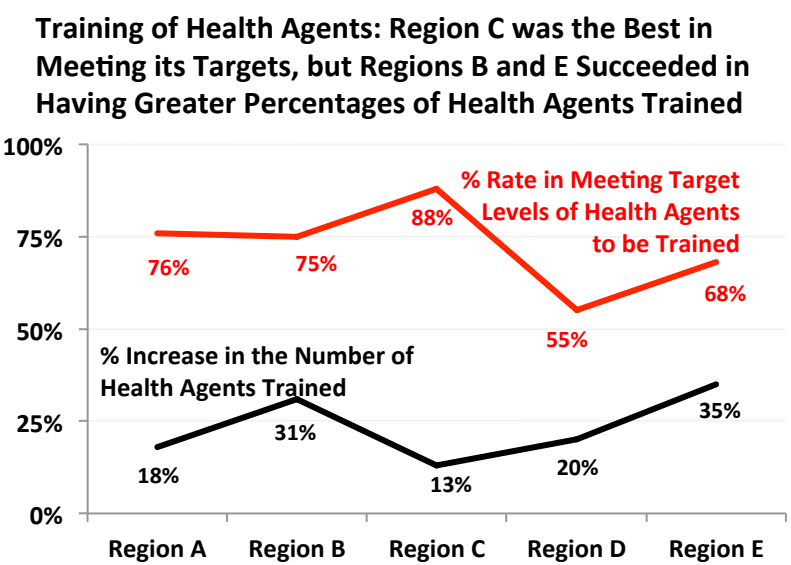


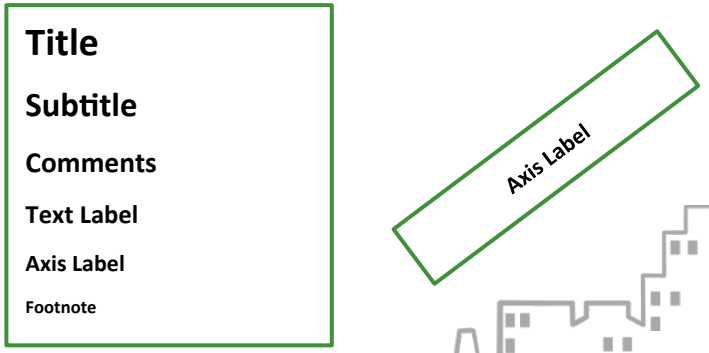
MLE

Four Principles:

Careful Choice of Pre-Attentive Attributes – **Positioning** (1)

- Ensure that the title is positioned in the upper left corner with left-side justification
- Add subtitles and comment boxes as needed to provide additional information
- Format the text within a data graphic to always run in a horizontal direction, with the exception that axis labels may run diagonally



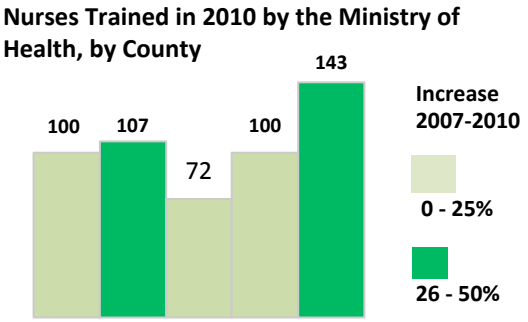


MLE

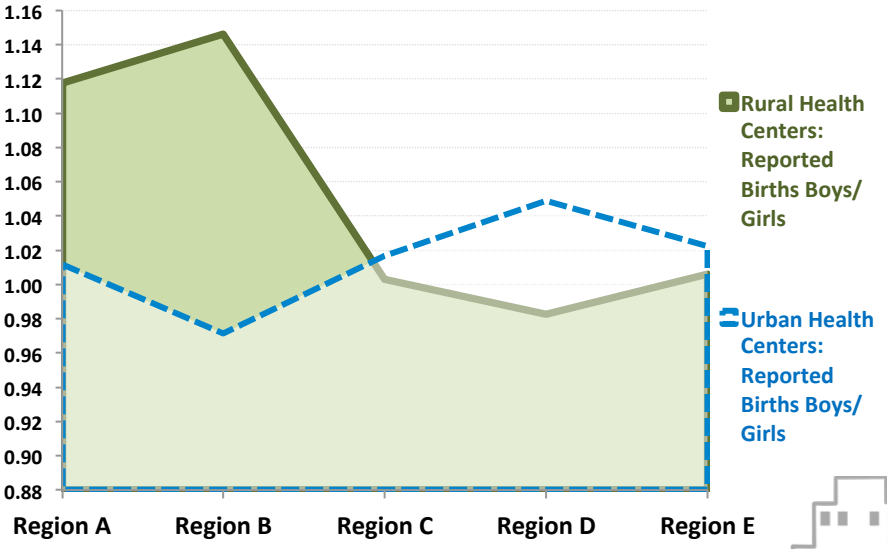
Four Principles:

Careful Choice of Pre-Attentive Attributes – **Positioning** (2)

- Place labels directly on the data elements in a data graphic, but use them prudently



- Consider thoroughly all the graphical means to better transmit the key message to be remembered
  - arrangement of elements
  - as well, the type of chart to use, text, colors and lines

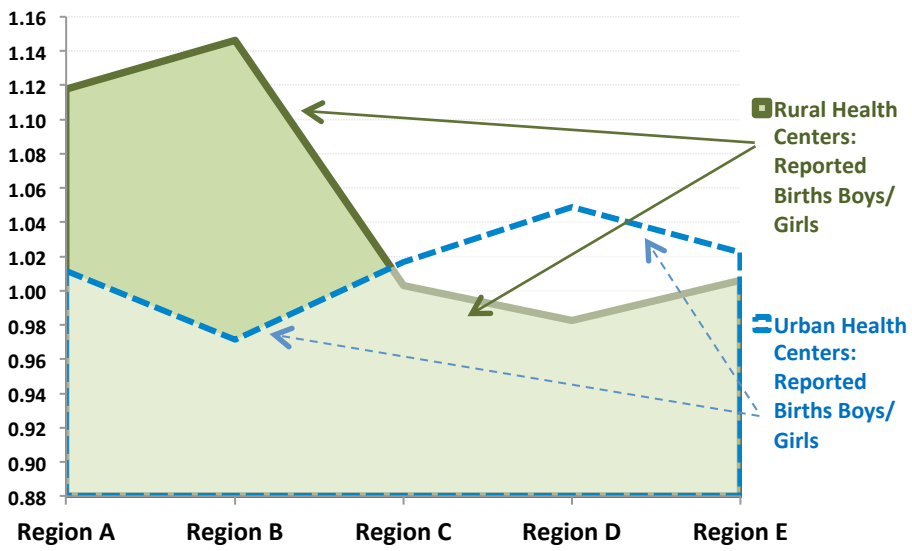


MLE

Four Principles:

Careful Choice of Pre-Attentive Attributes – **Movement** (1)

- Use wisely the movement of elements in a data graphic
  - Dynamic movement of elements can as easily attract a viewer's attention as distract it



MLE

Four Principles related to the Development of a Data Visualization

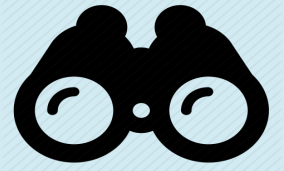
- Selection of the Appropriate Graphic
- Simplicity of Style in Presentation
- Careful Choice of Pre-Attentive Attributes
- **Honesty**

MLE

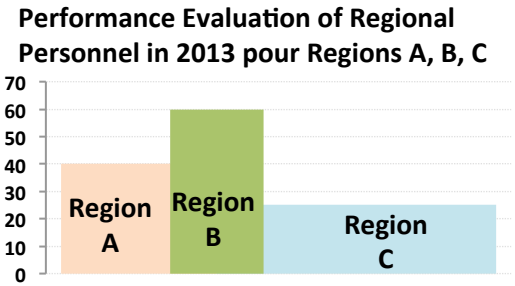
Four Principles :

Evaluating the Presentation of Data for Honesty – Considerations for Accuracy

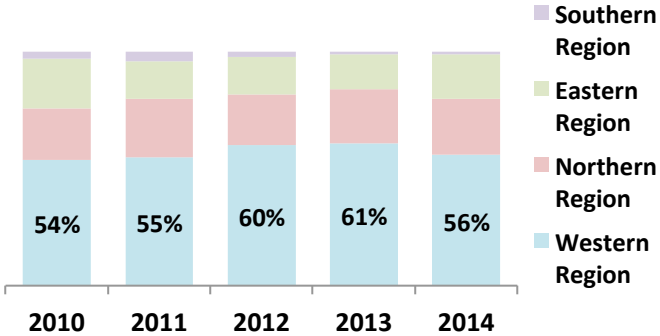
- First start off by verifying the quality of the data



- Always ensure that spaces between intervals in a data graphic are the same width along their axis



- Ensure that percentages of wedges in a pie chart add up to 100% (similarly the case for other data graphics whose component parts of a whole are presented as percentages)

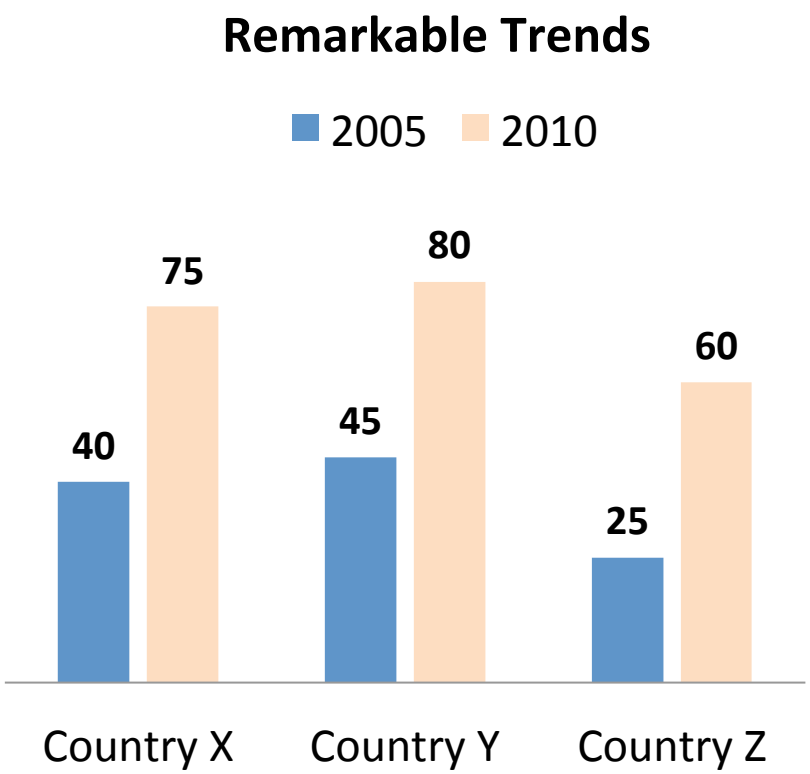


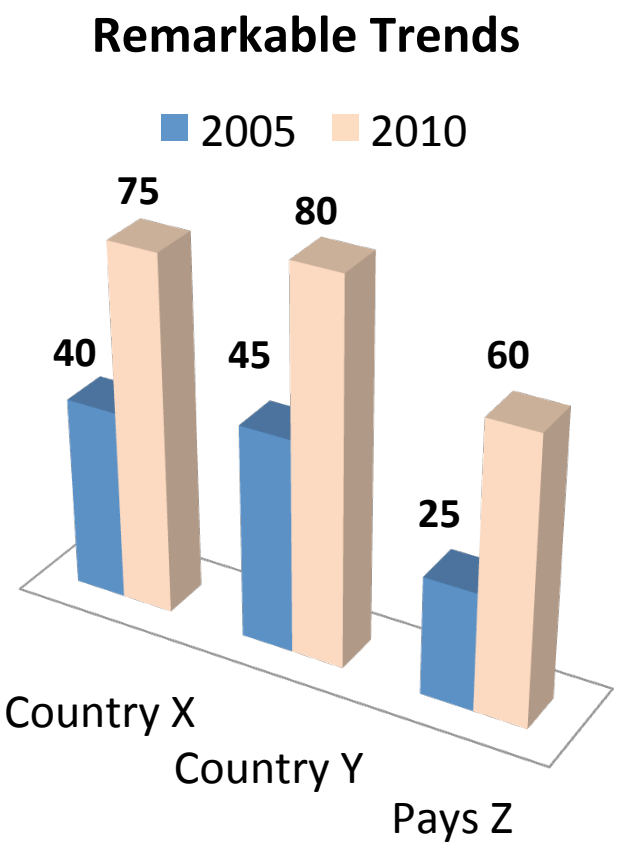
MLE

Four Principles :

Evaluating the Presentation of Data for Honesty –

Avoid the Use of Three-Dimensional Graphics (1)



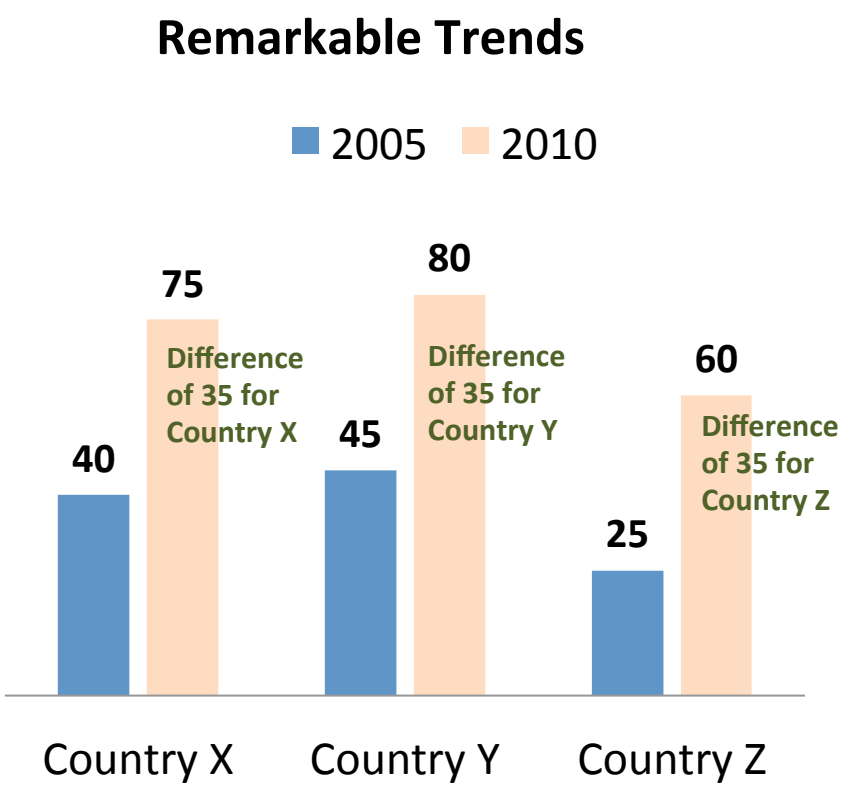


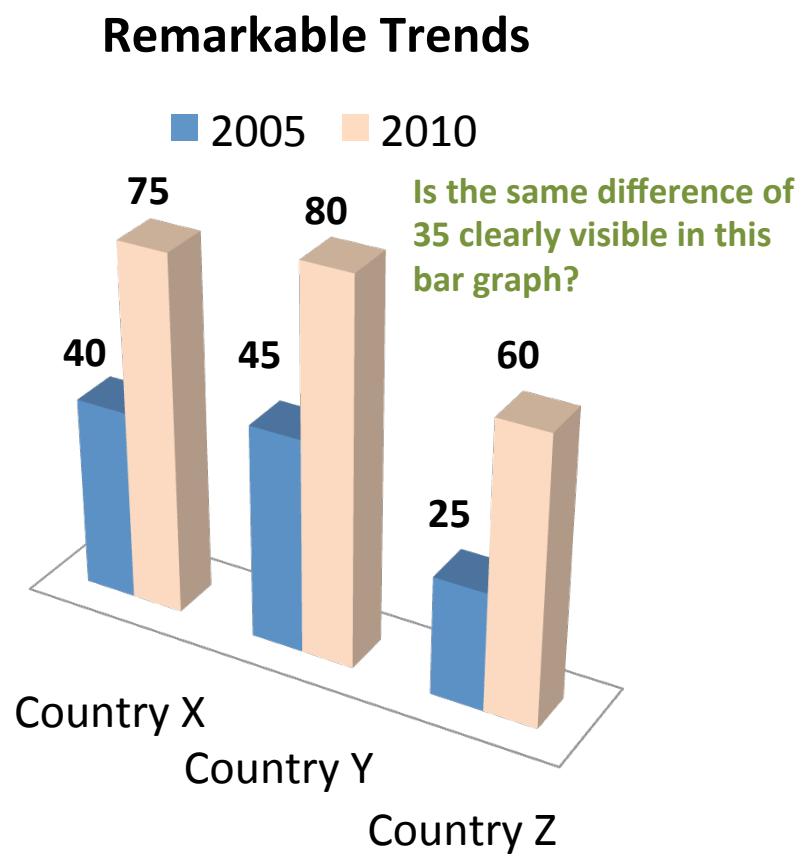
MLE

Four Principles :

Evaluating the Presentation of Data for Honesty –

Avoid the Use of Three-Dimensional Graphics (2)



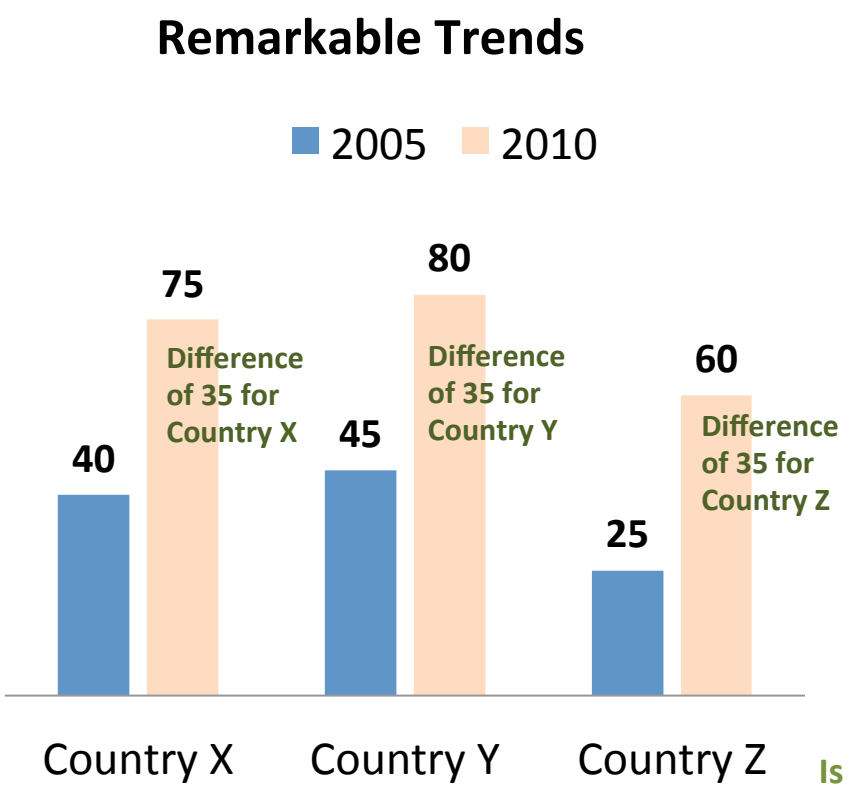


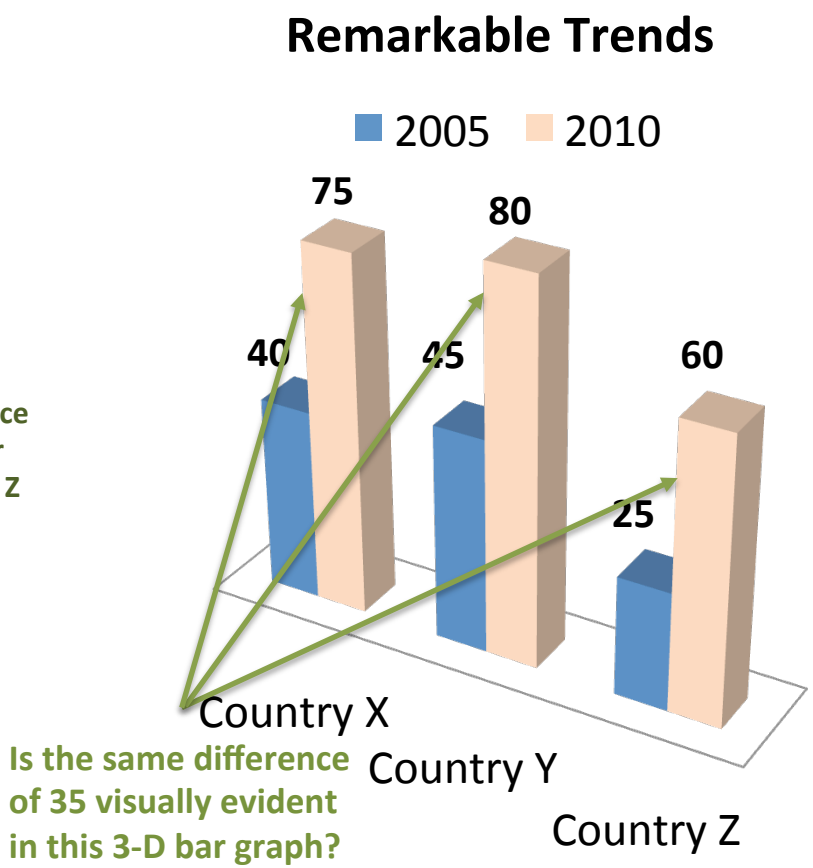
MLE

Four Principles :

Evaluating the Presentation of Data for Honesty –

Avoid the Use of Three-Dimensional Graphics (3)





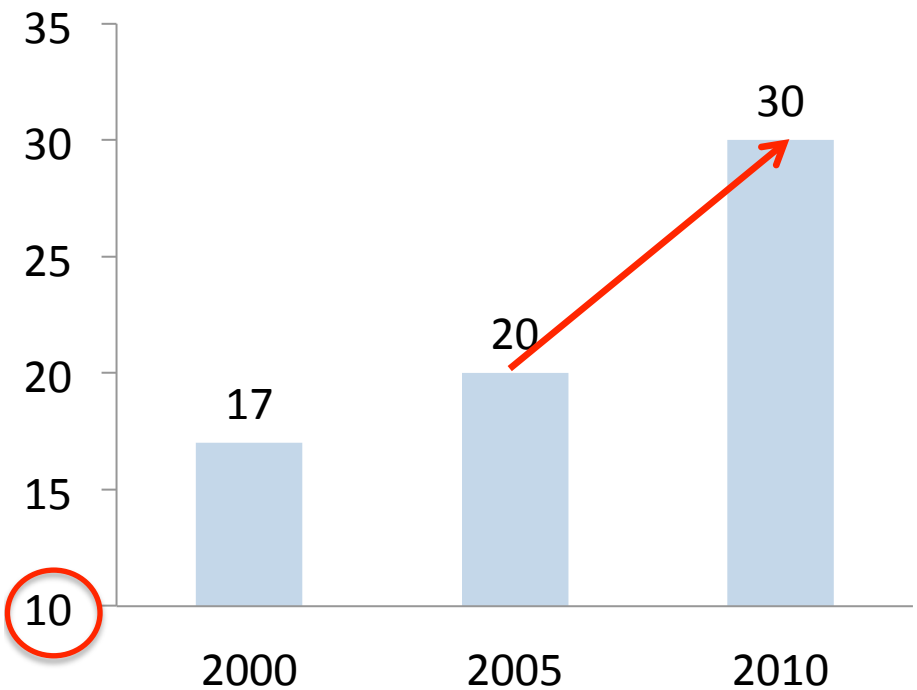
MLE

Four Principles :

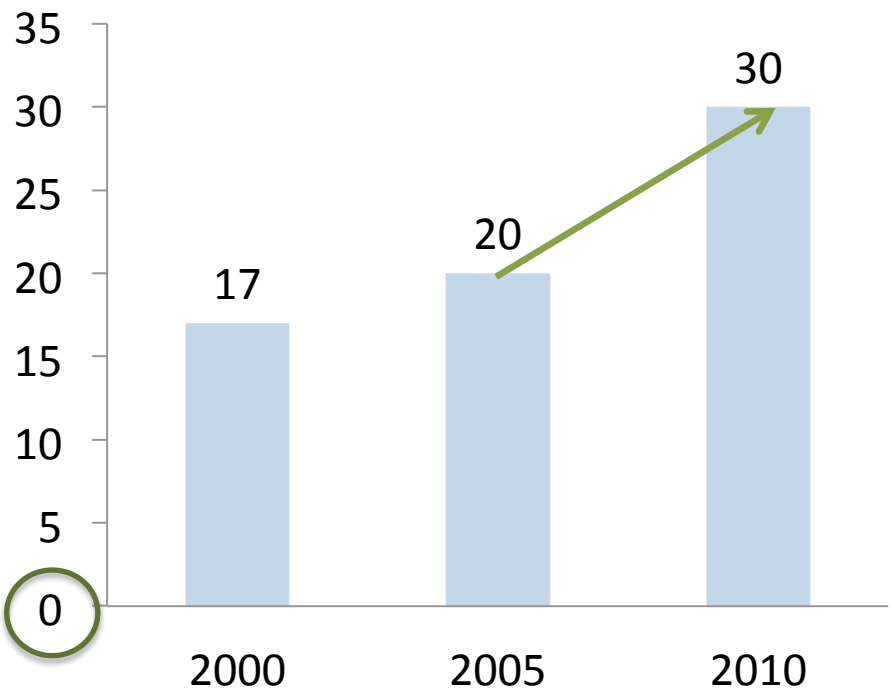
Evaluating the Presentation of Data for Honesty –

Axes of Quantitative Values Must Begin at Zero (1)

Observed Trend in Country X
on the Use of
Modern Family Planning Methods



Observed Trend in Country X
on the Use of
Modern Family Planning Methods

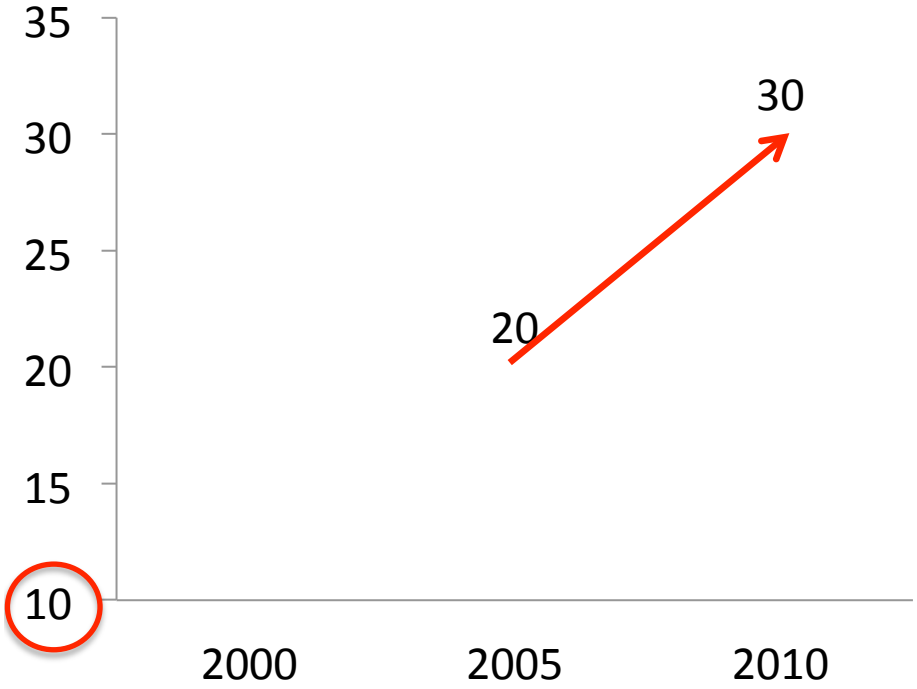


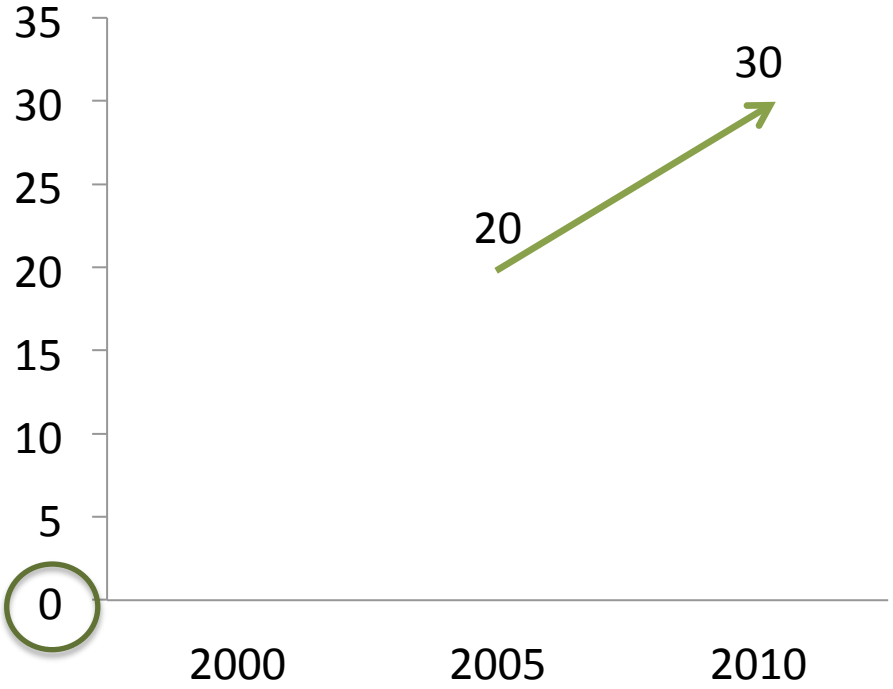
MLE

Four Principles :

Evaluating the Presentation of Data for Honesty –

Axes of Quantitative Values Must Begin at Zero (2)



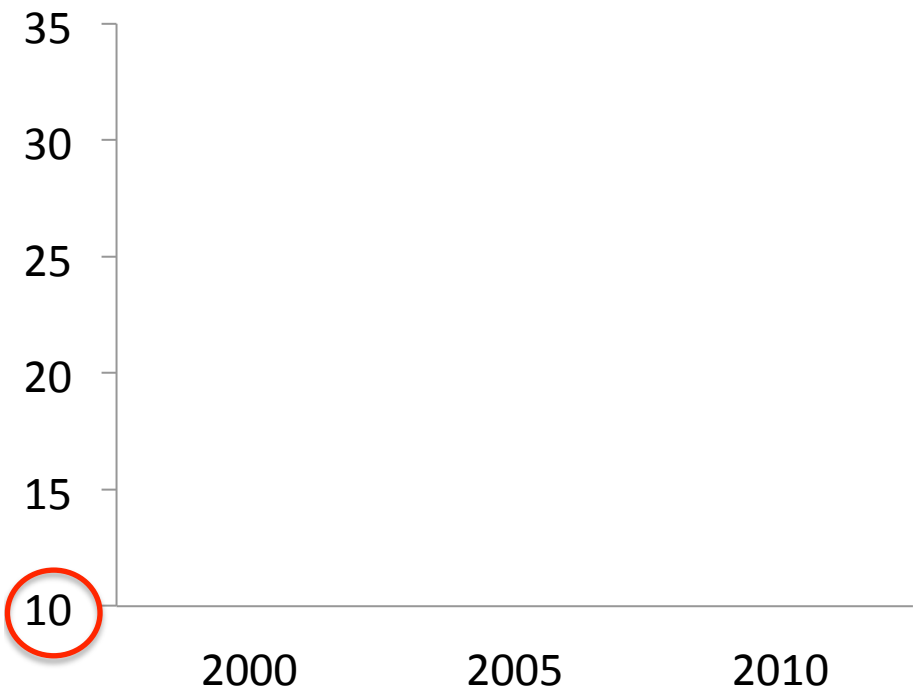


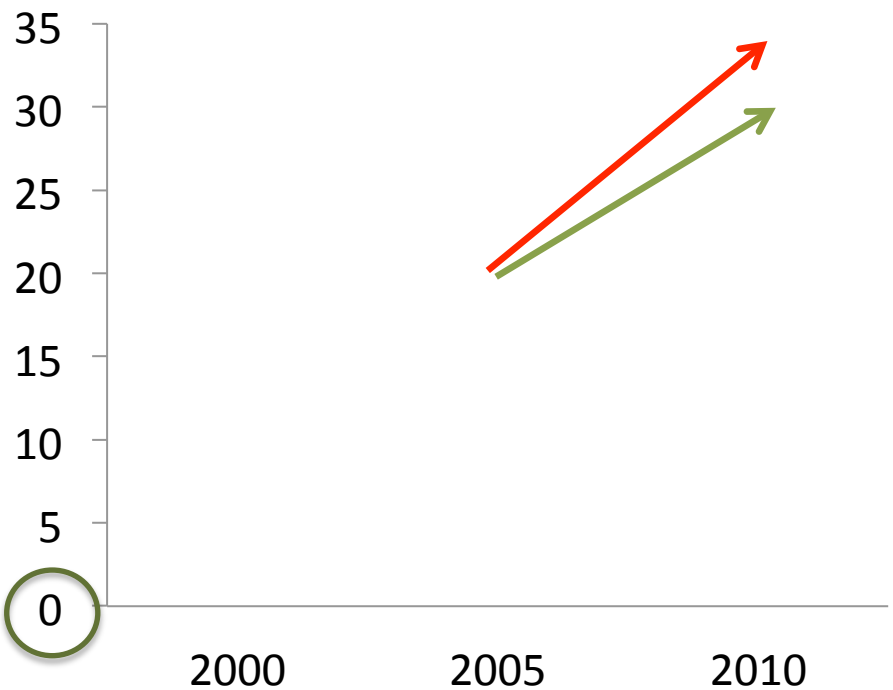
Four Principles :

Evaluating the Presentation of Data for Honesty –

Axes of Quantitative Values Must Begin at Zero (3)

The slope angle of the arrow is greater – falsely indicating a larger rate of change over time – for the case when the vertical axis begins at 10 rather than at 0





Coffee / Tea Break (15 minutes)

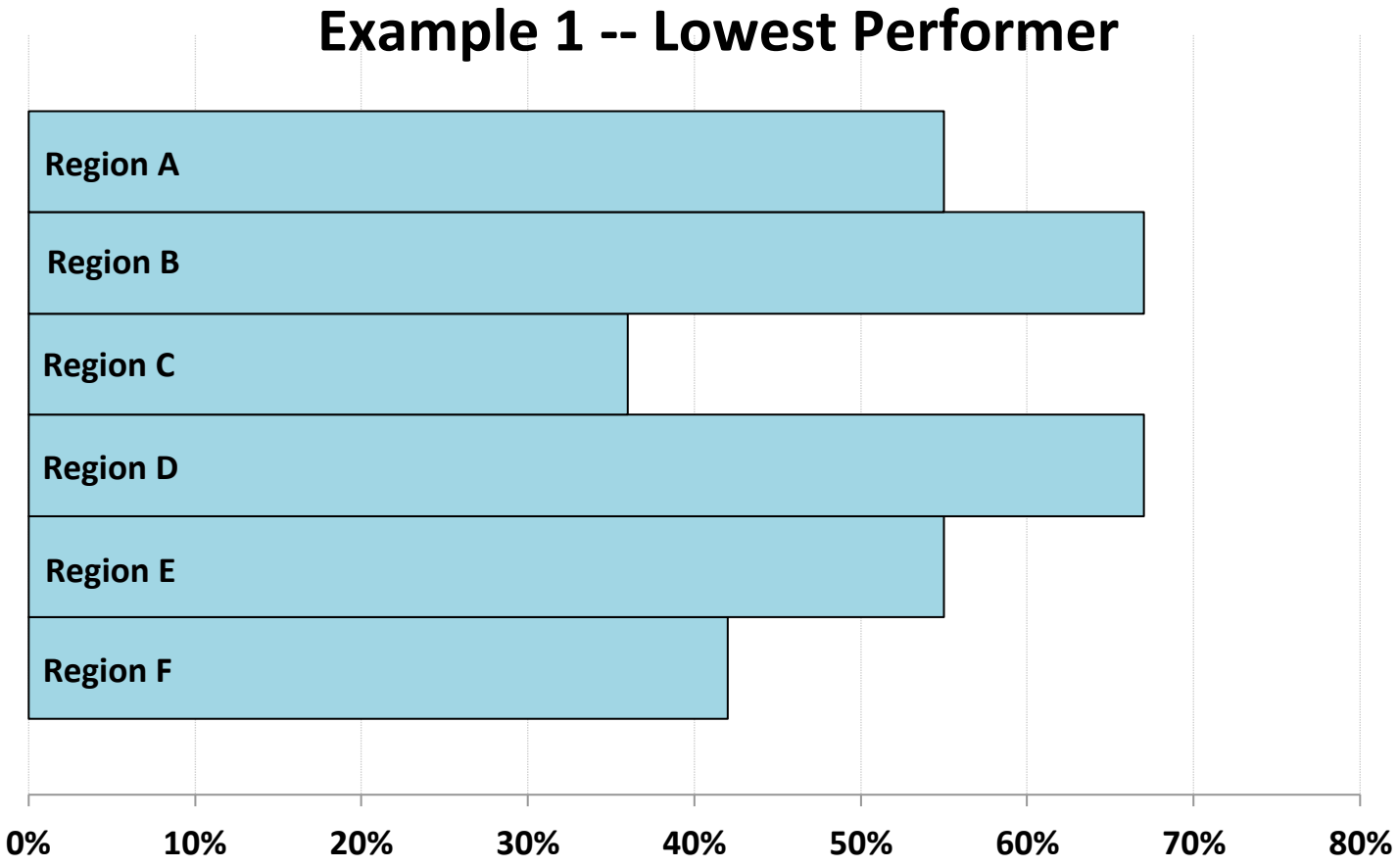
Interactive Activity #2 (20 minutes)

Remaining in plenary:

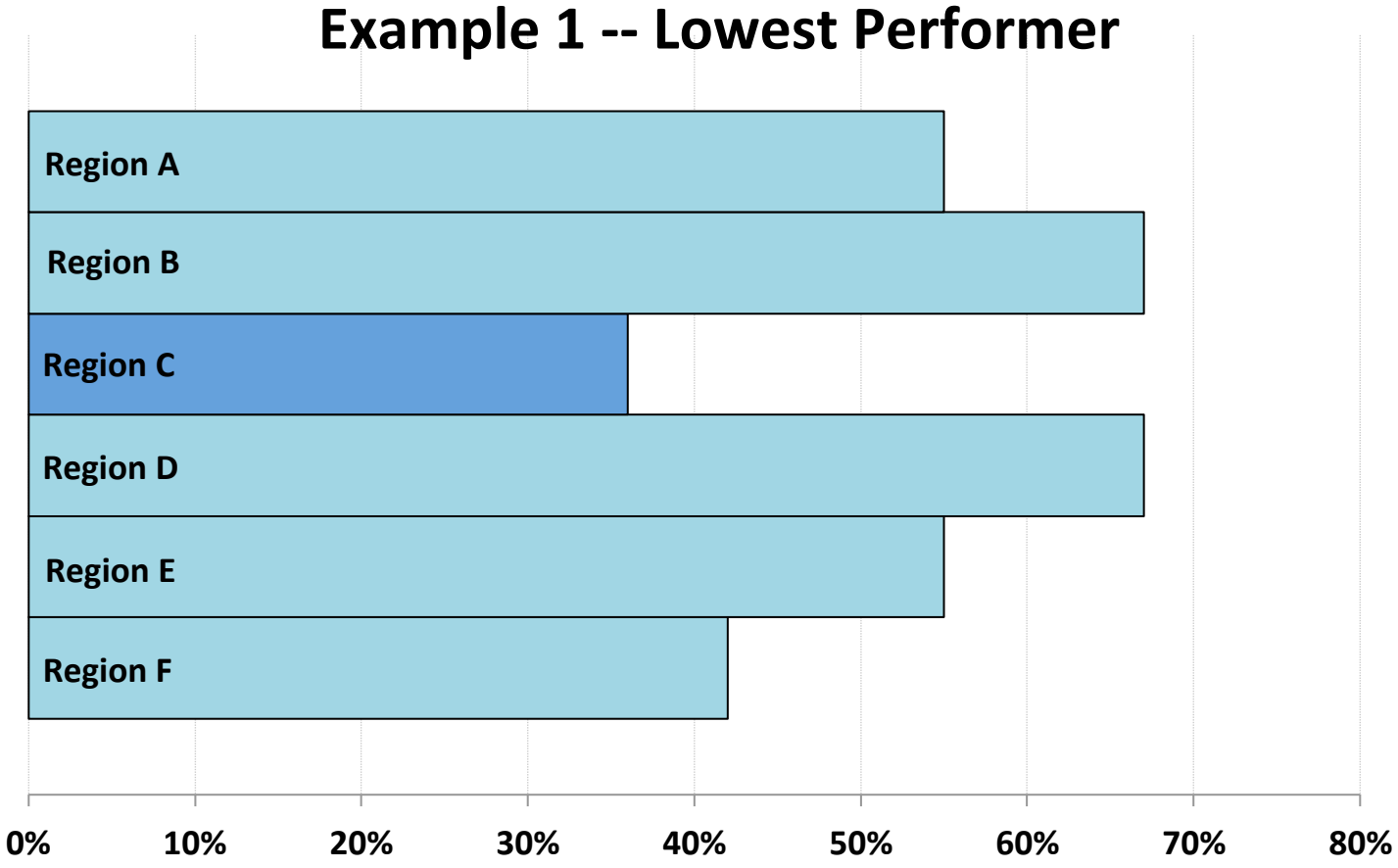
- Evaluate and revise these examples of corrected « bad » data graphics, based on what you have learned from the previous session
- Formulate a title for each corrected data graphic so that it gives an accurate summary of the findings from the data analysis

MLE

Interactive Activity #2 (continued)



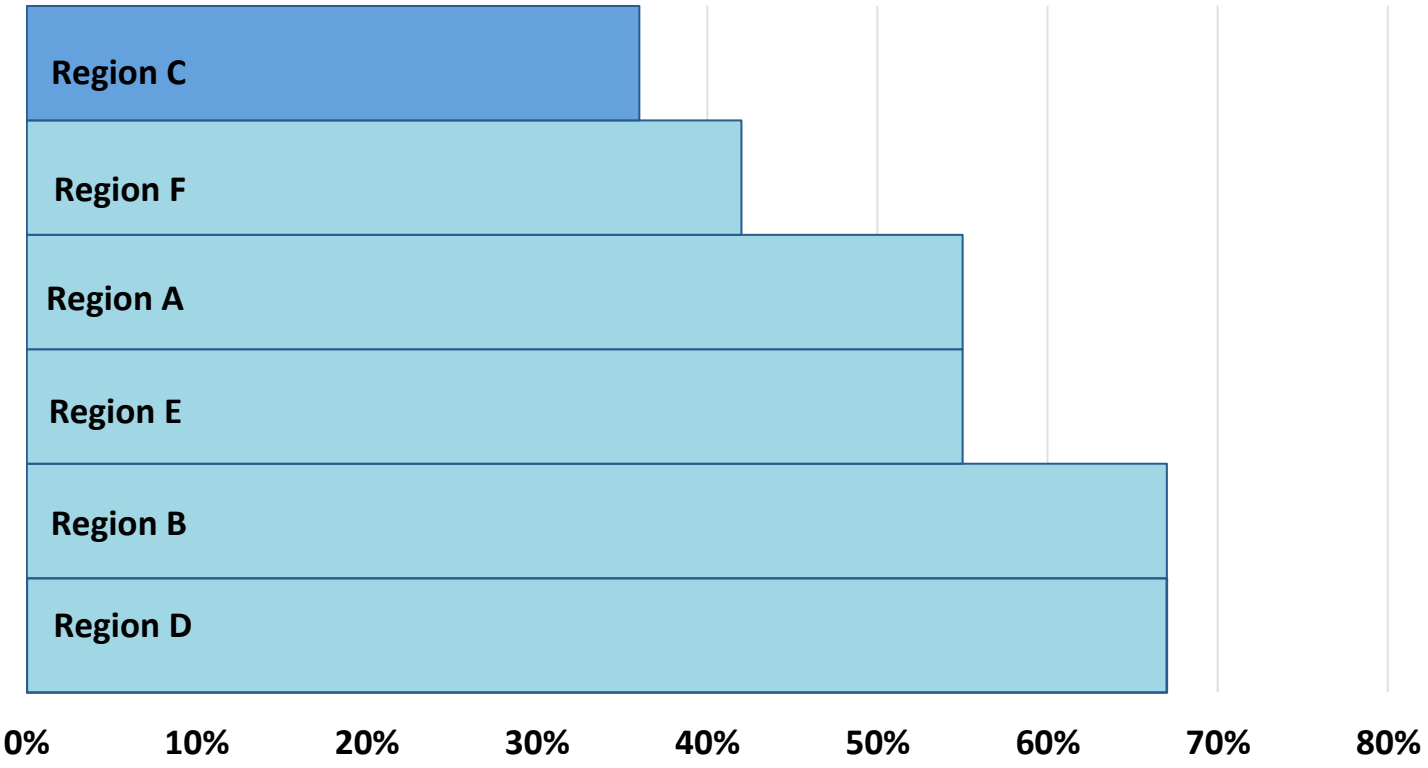
Interactive Activity #2 (continued)



MLE

Interactive Activity #2 (continued)

Example 1 -- Lowest Performer

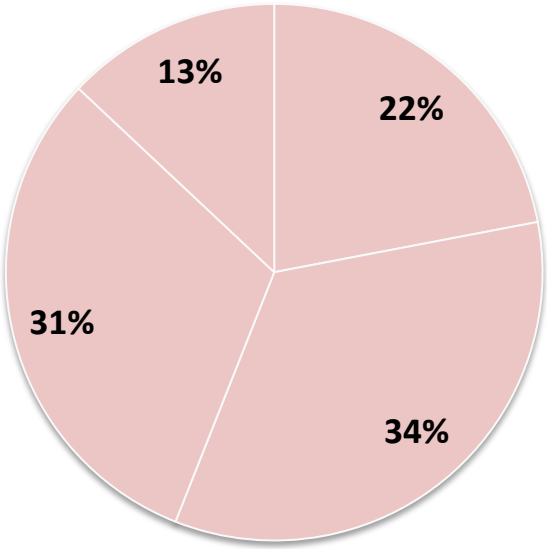


MLE

Interactive Activity #2 (continued)

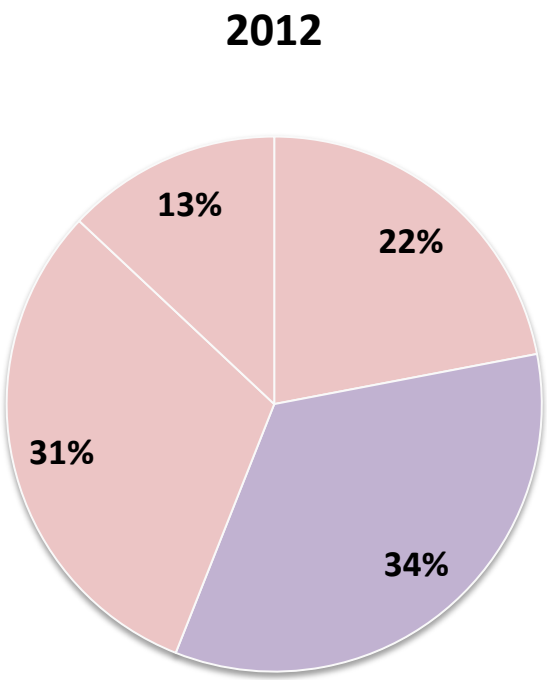
Example 2

2012



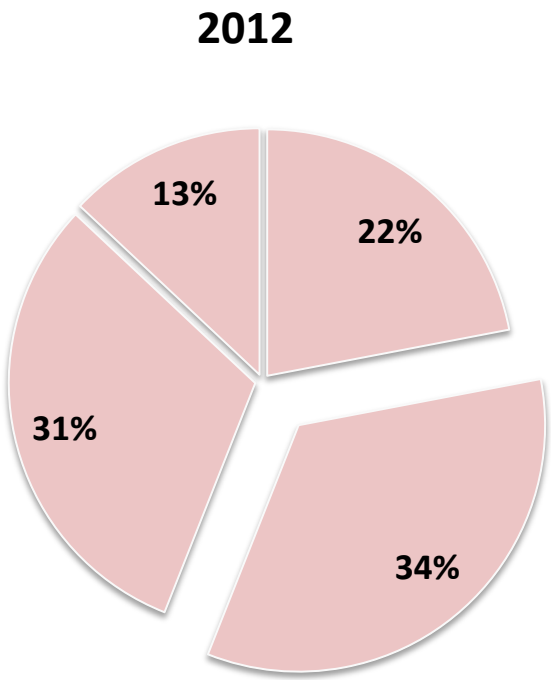
Interactive Activity #2 (continued)

Example 2 – Improved version



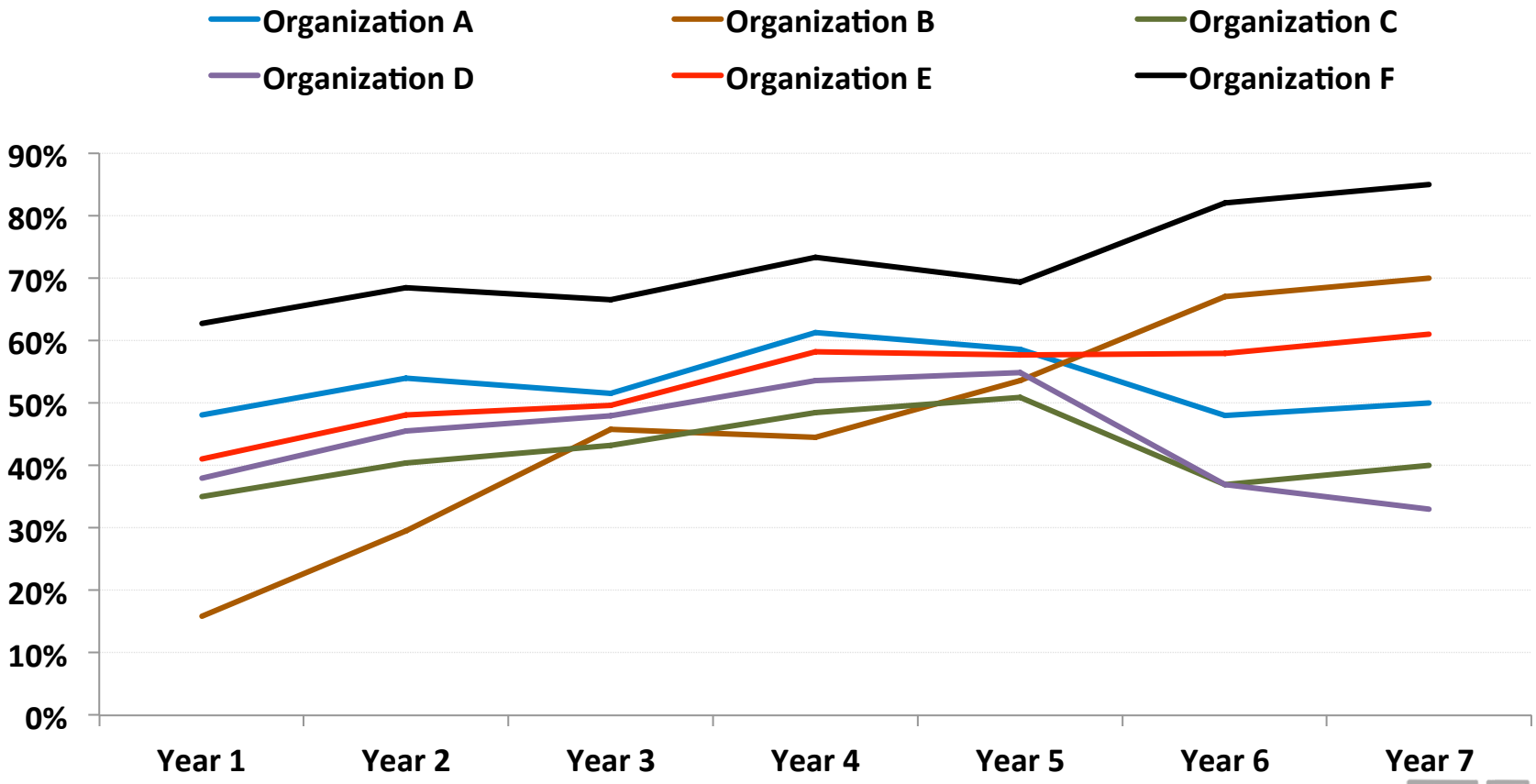
Interactive Activity #2 (continued)

Example 2 – Another improved version



Interactive Activity #2 (continued)

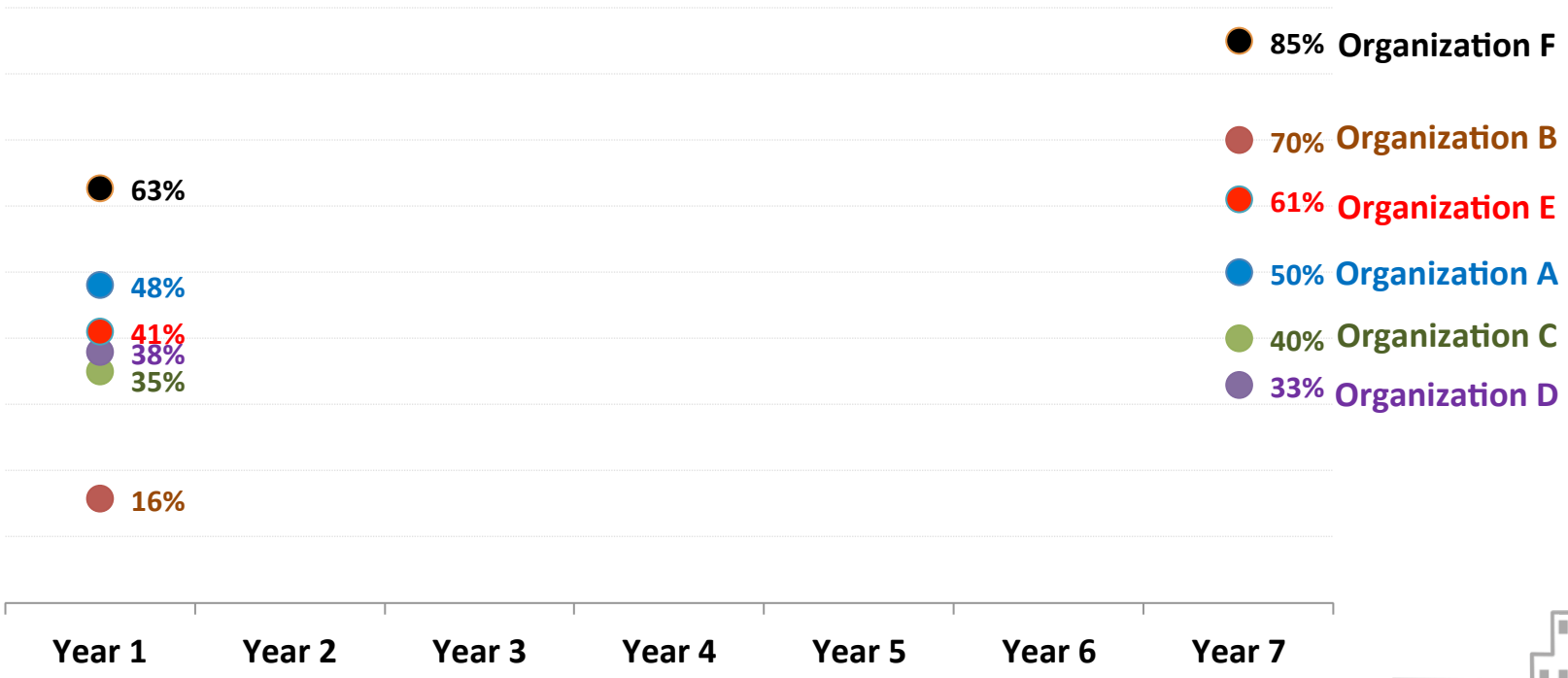
Example 3



MLE

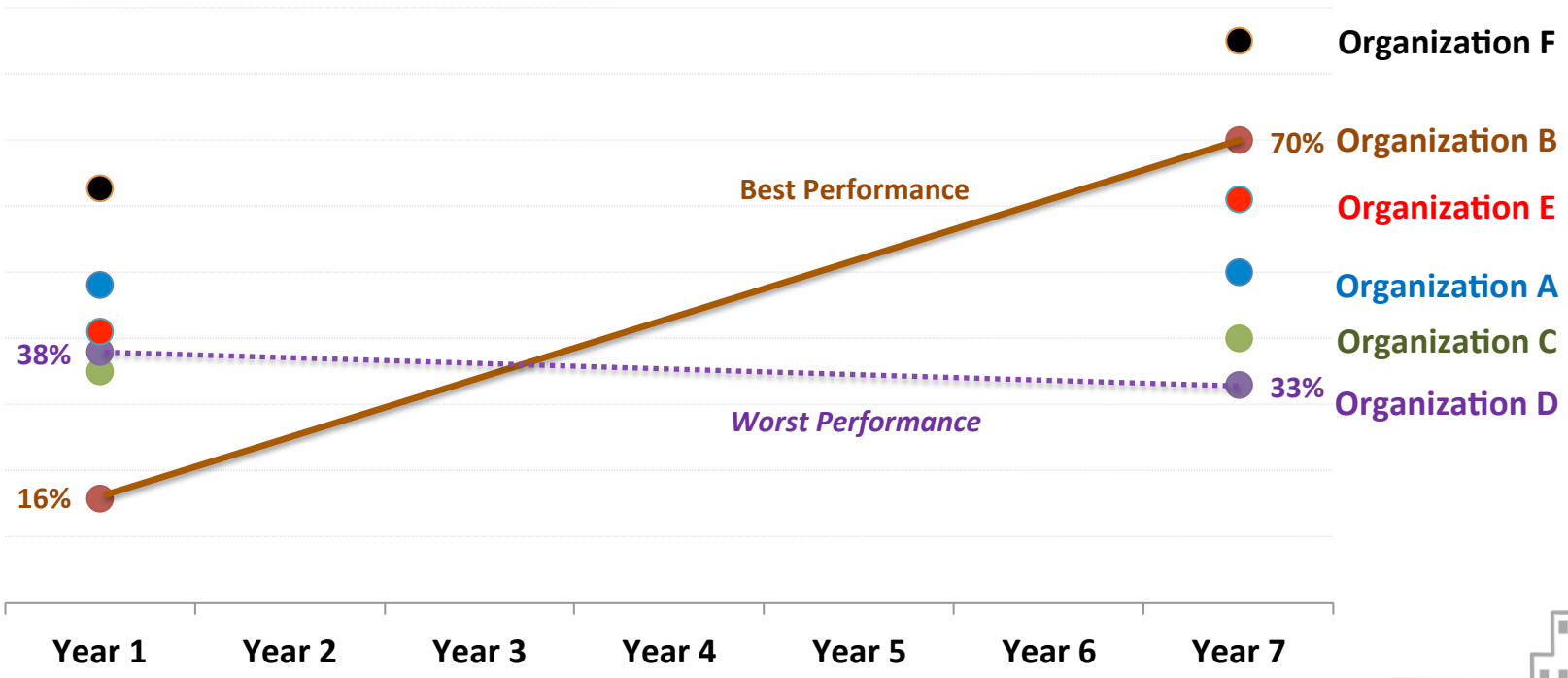
Interactive Activity #2 (continued)

Example 3 – Improved version



Interactive Activity #2 (continued)

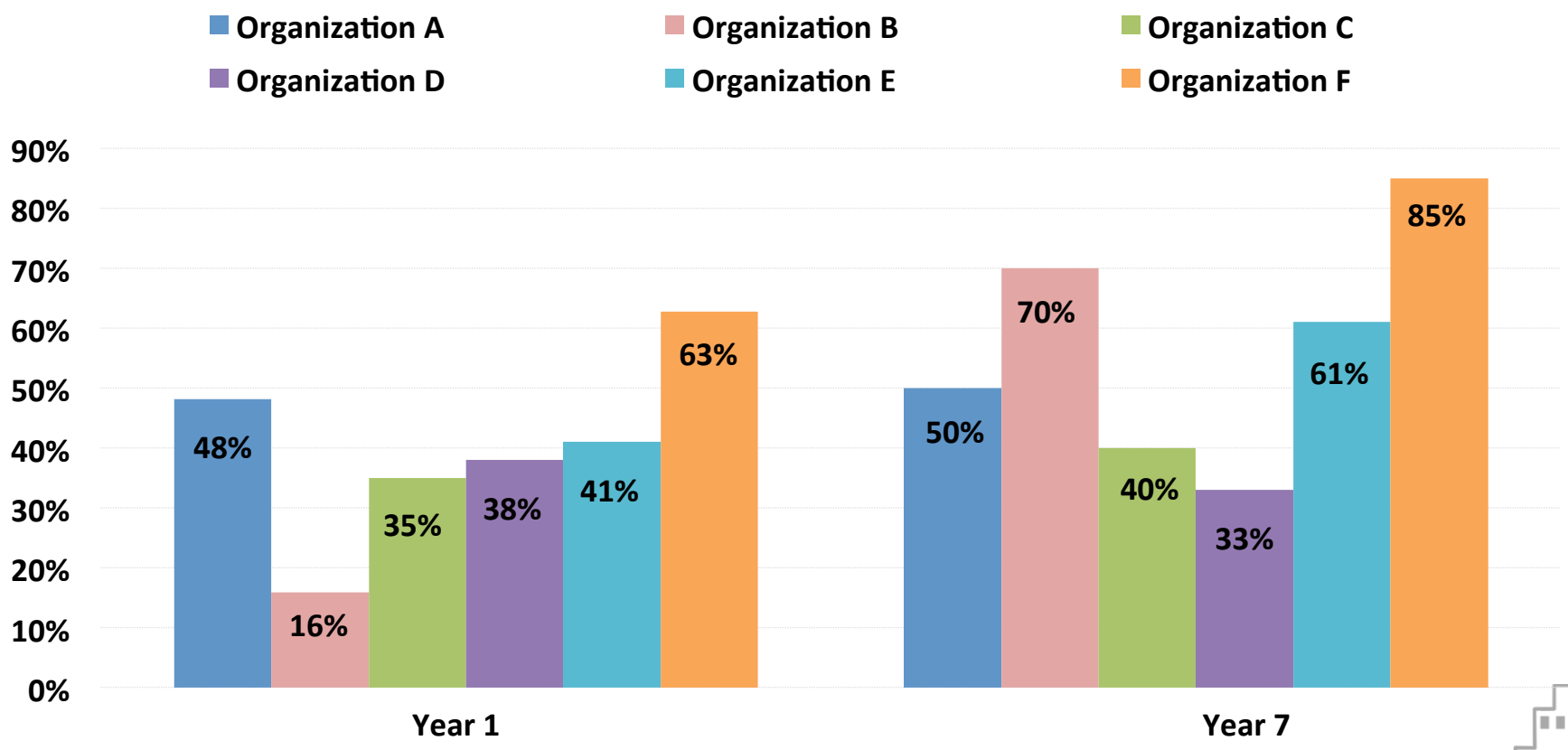
Example 3 – Another improved version



MLE

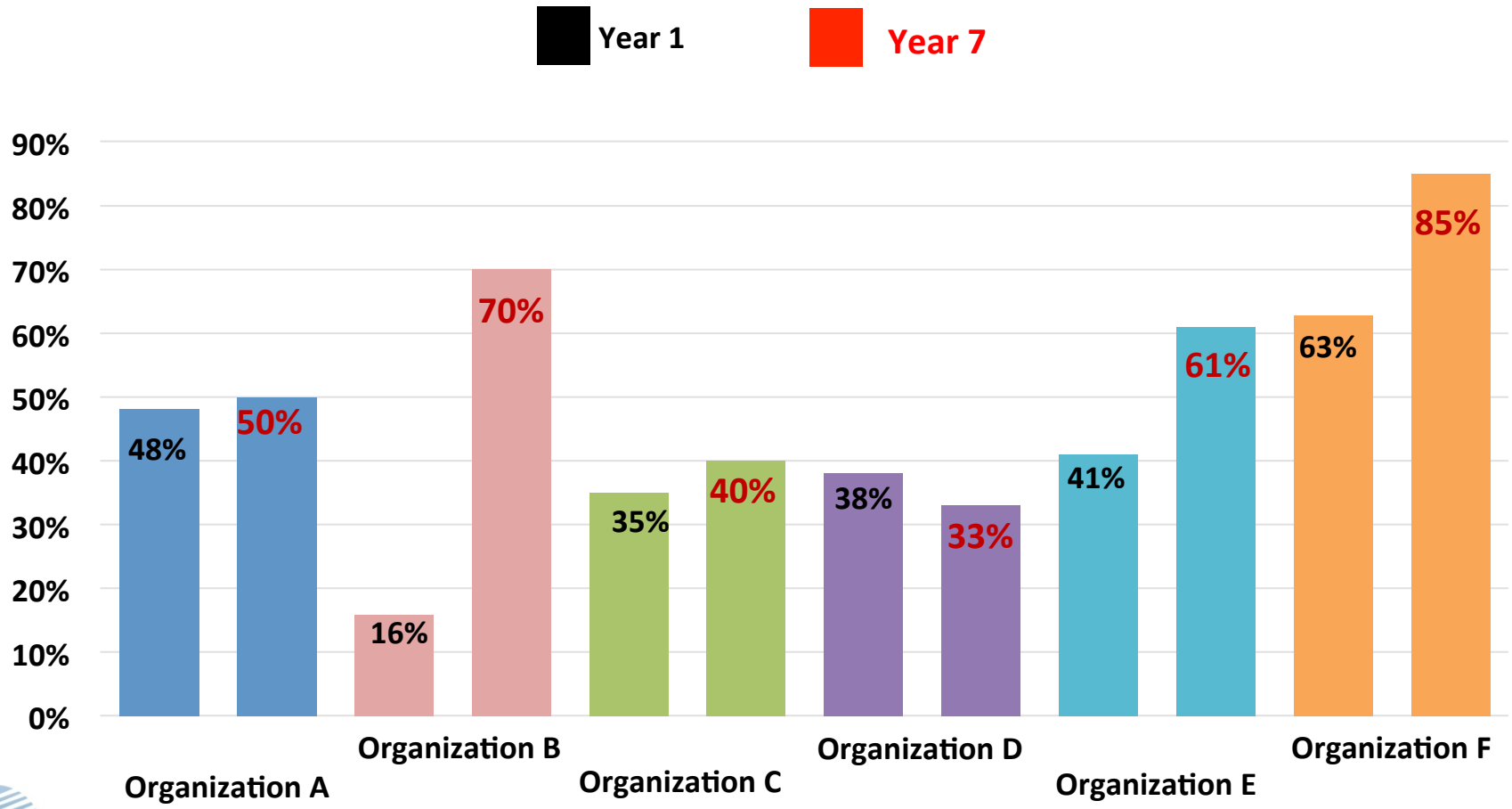
Interactive Activity #2 (continued)

Example 4



Interactive Activity #2 (continued)

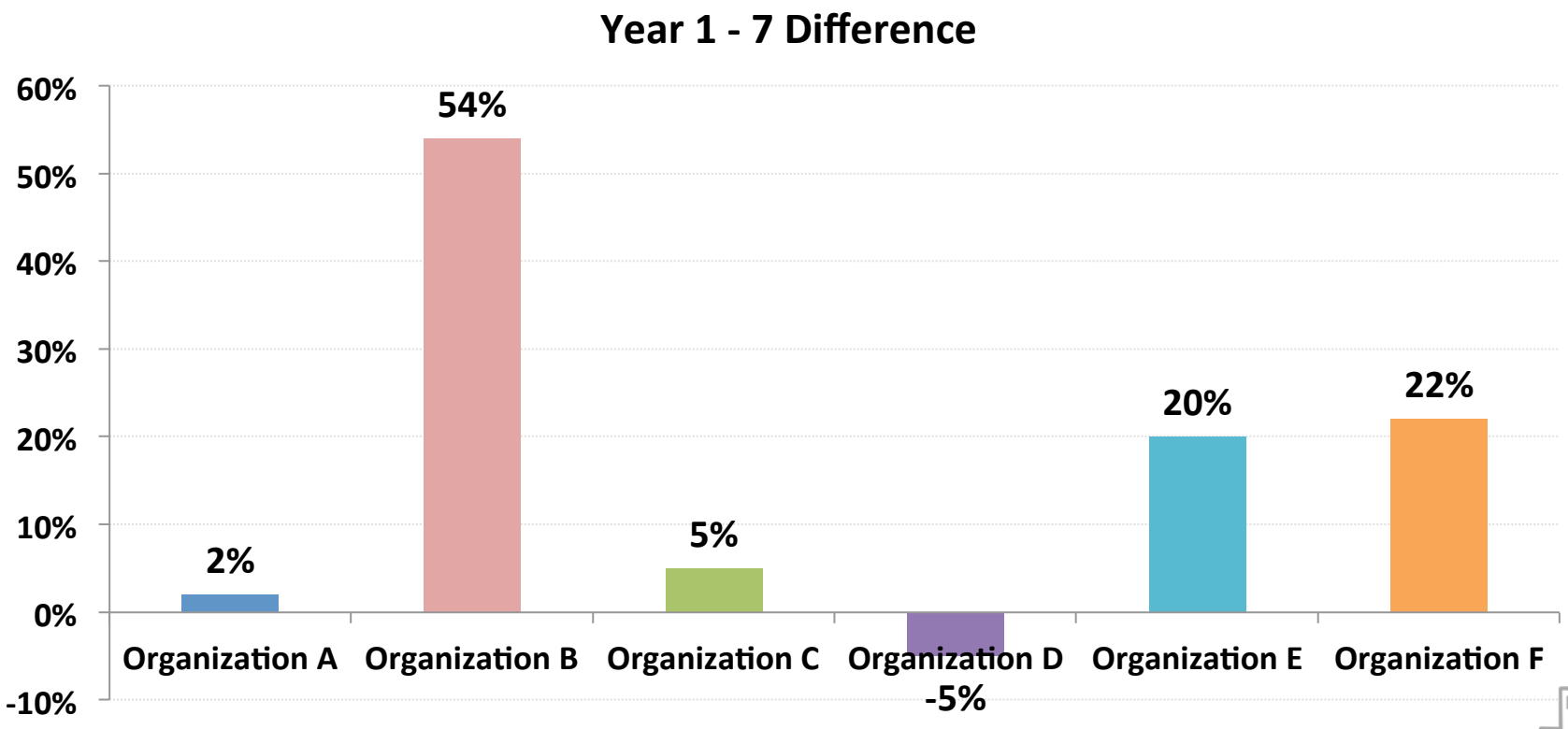
Example 4 – Improved version



MLE

Interactive Activity #2 (continued)

Example 4 – Another improved version

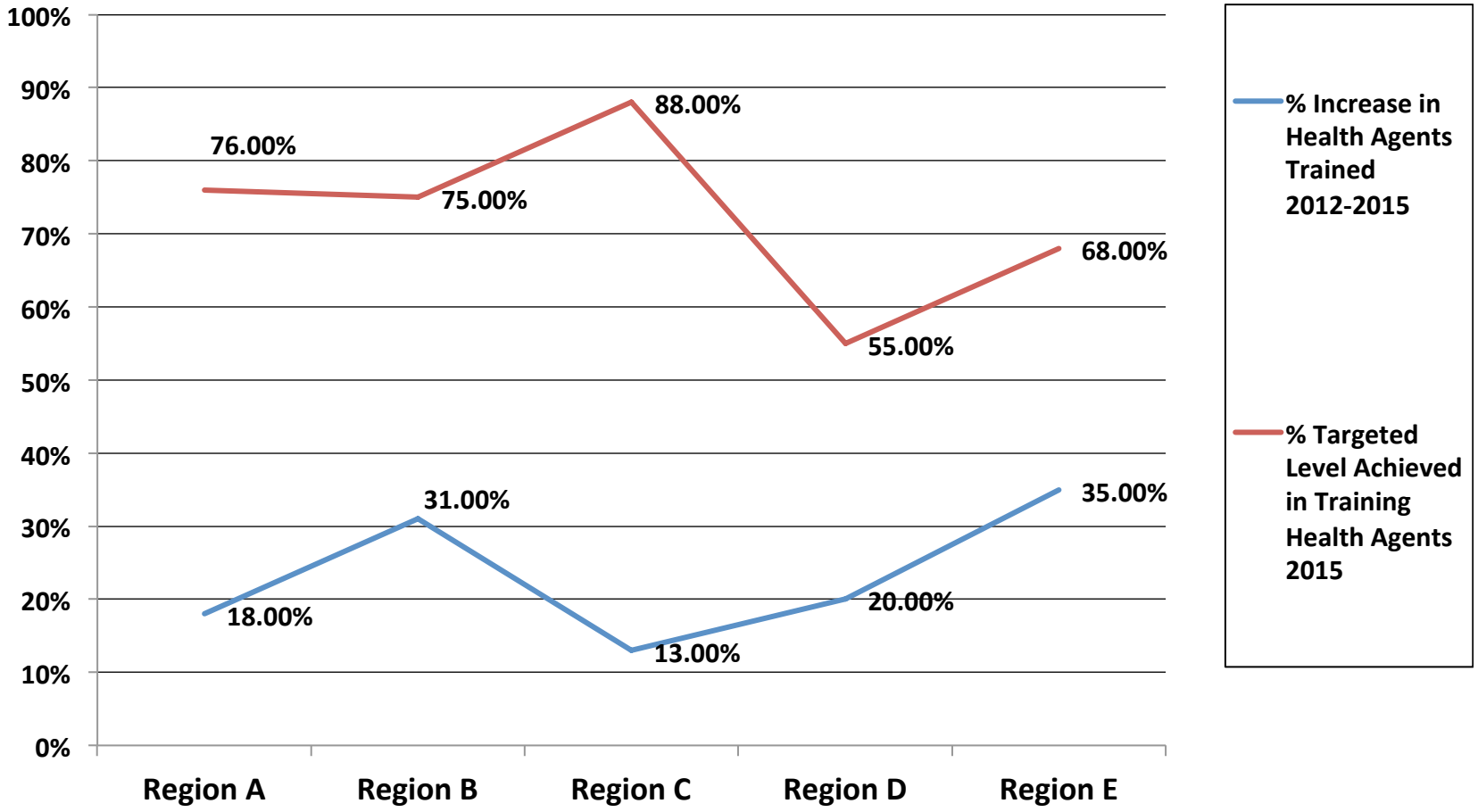


MLE

Interactive Activity #3 (55 minutes)

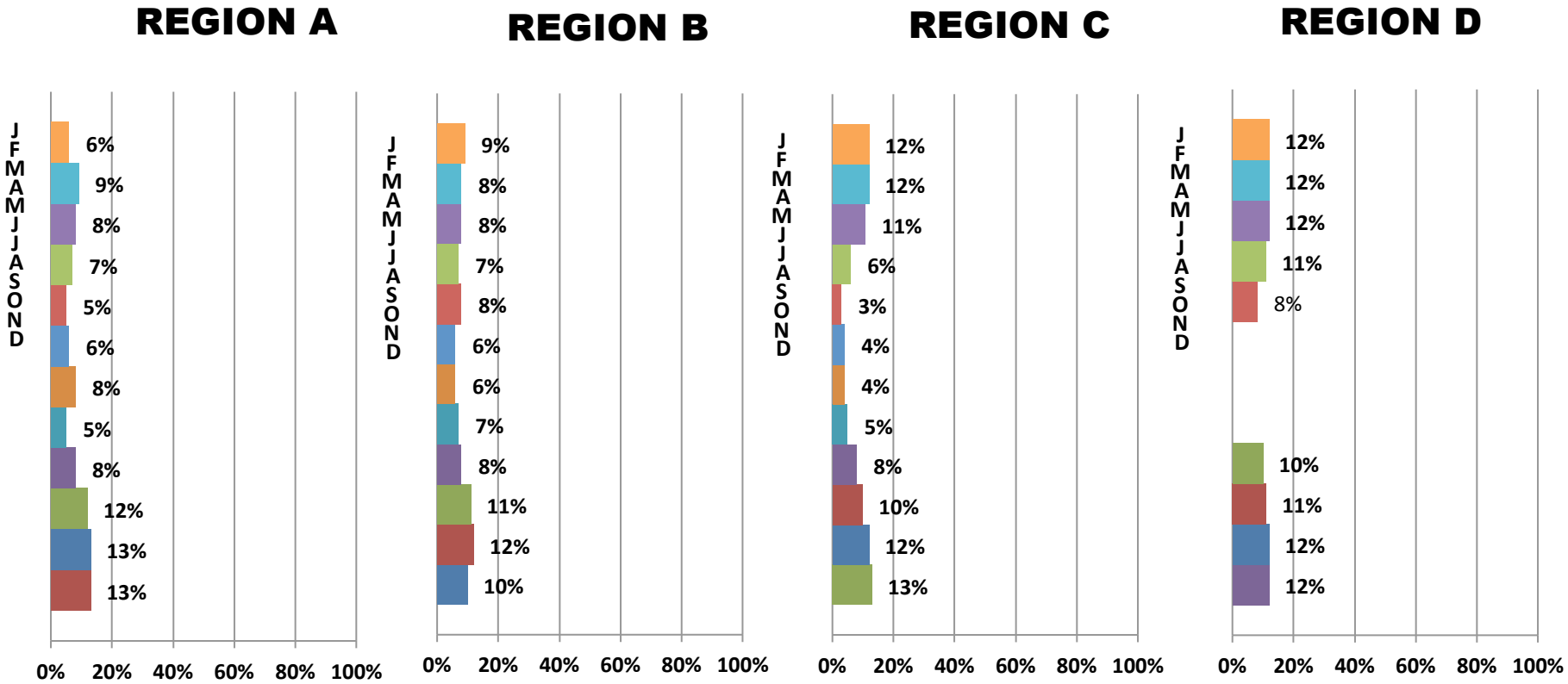
- Separate yourselves into pairs/groups in order to critique each of the following data visualization(s), as based on what you have learned about the four principles in developing a data visualization
- List out the various problems you identify with each data visualization
- Describe how you would revise each data visualization to make it more effective

Training of Health Agents in 2012 and 2015



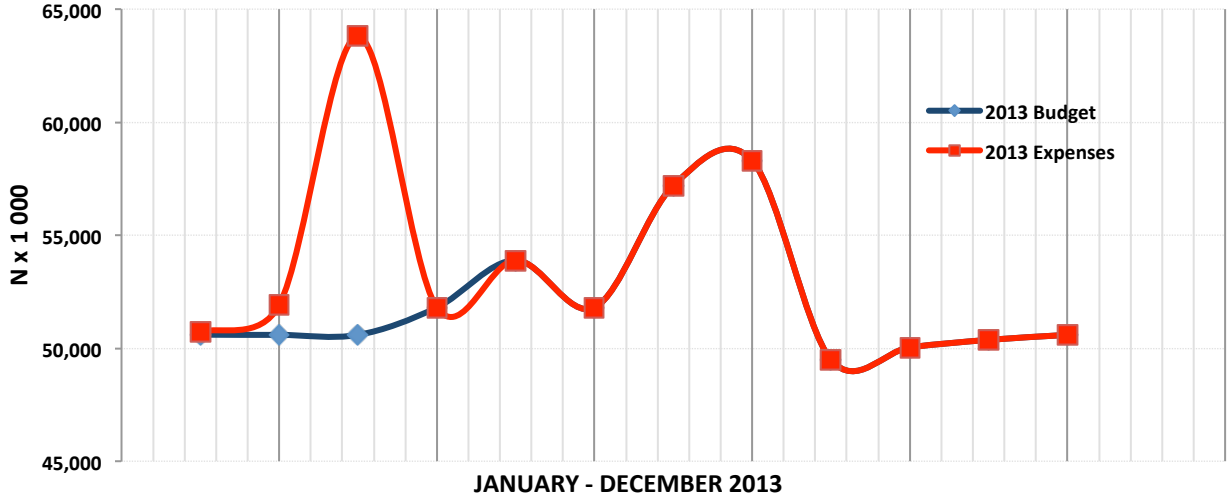
MLE

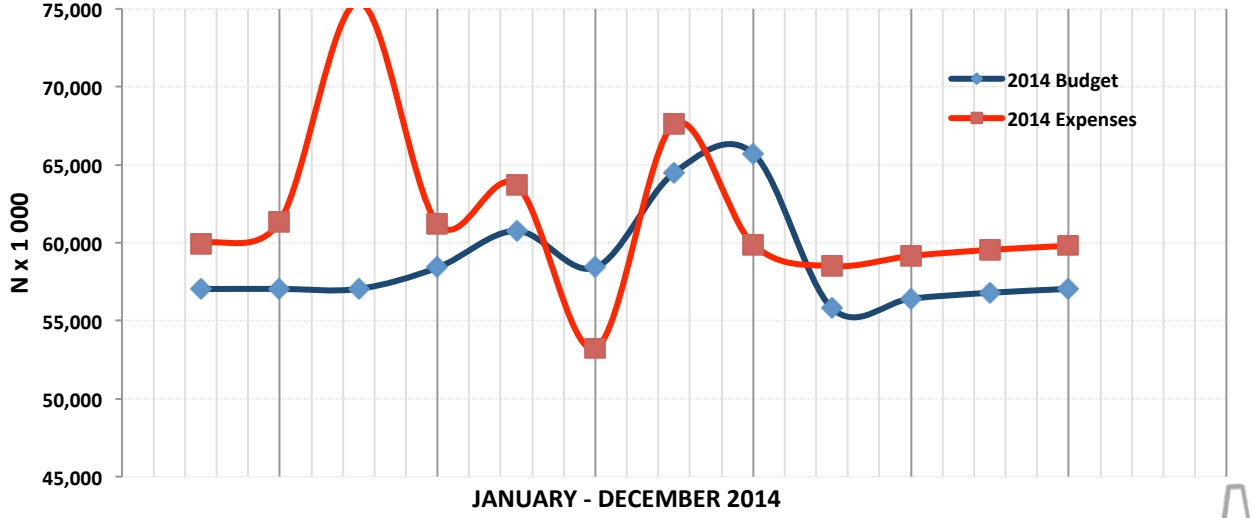
PERCENTAGE OF ANNUAL OUTPATIENT VISITS CUMULATIVELY SEEN PER MONTH FOR 2015 BY HEALTH POSTS WITHIN EACH REGION -- REGIONS A, B, C & D



MLE

2013 & 2014 MONTHLY OPERATIONAL BUDGETS AND EXPENSES FOR REGION X





MLE

Developing a Data Visualization:
Applying What You Have Learned

As objectives, at the end of this session, participants will be able to:

- Identify one or more key messages to transmit based on research data and formulated for a specific audience
- Develop one or more data visualizations that communicate follows the four principles of a data visualization and that clearly conveys the expected message to the targeted audience

(30 minutes)

MLE

Interactive Activity #4 (15 minutes)

- By means of a brainstorming session, a brief recap is done on all the topics covered throughout this training until now under three chapters :

Interactive Activity #4 (15 minutes)

- By means of a brainstorming session, a brief recap is done on all the topics covered throughout this training until now under three chapters :
  - Identification of a targeted audience
  - Identification of a message based on the data to transmit
    - patterns, trends, comparisons, and outliers
    - method for storyboarding
  - Development of a data graphic
    - scientific aspects of visual perception
    - four principles of data visualizations
- These topics are listed out on paper, which is attached onto a wall for reference by the participants

MLE

Interactive Activity #5 (120 minutes)

- Separate yourselves into pairs/groups
- Get assigned a targeted audience (e.g., decision maker, donor, journalist)
- Analyse a table of research data given to you in order to formulate one or more key findings that your pair/ group wants to share with that targeted audience
- Develop at least one data visualization based on those key findings for the targeted audience assigned to you

MLE

Interactive Activity #5 (continued)

(120 minutes)

- Provide the reasons by which you have developed your data visualizations, noting on sheets of poster paper:
  - Your analysis of the data and the key finding that you have extracted from the data by means of storyboarding
  - Scientific aspects of perception that you considered in the design of your data visualizations
  - Principles of data visualization that you considered in the design

MLE

Interactive Activity #5 (continued)

(120 minutes)

- Exchange your data visualizations with those of another pair/group who had been assigned a different target audience
- Explain on sheets of poster paper the reasons by which you believe the other pair/group had developed their data visualizations
- Place the two sets of poster papers for each set of data visualizations side by side for comments by the plenary

MLE

Percentage Distribution with Respect to Economic Level and City Residence of Women (15 – 49 years of age) Currently using a Method of Contraception: MLE Project Baseline Study (2010) and End-line Study (2014) -- Nairobi, KENYA

	ANY METHOD	ANY MODERN METHOD	Sterili-zation	Implnts	IUD	Inject-ables	Pills	Emergncy Contra-ception	Male Con-doms	Stand. Days Methd	Other Modern Methd	ANY TRADITNAL METHOD	*NO METHOD*	N =
NAIROBI at Baseline														
Economic Level														
Lowest Level	41.7	37.6	1.7	1.1	1.2	21.9	5.3	1.0	4.9	--	0.6	4.2	*58.3*	483
Second Level	47.9	44.6	1.6	1.4	1.9	25.6	7.9	1.0	4.0	--	1.3	3.3	*52.1*	518
Middle Level	57.2	53.8	0.6	2.5	1.2	22.0	19.6	0.6	7.2	--	0.1	3.4	*42.8*	522
Second Highest	47.5	42.5	0.9	2.4	1.4	17.0	11.5	1.4	7.5	--	0.5	5.0	*52.5*	549
Highest Level	44.9	40.1	2.7	4.1	4.7	5.2	10.7	1.2	11.6	--	0.0	4.7	*55.1*	634
ALL LEVELS	47.8	43.6	1.5	2.4	2.2	17.7	11.1	1.0	7.2	--	0.5	4.2	*52.2*	2,706
SUBSET: WOMEN in a RELATNSHIP	63.0	58.0	2.2	3.4	3.1	26.9	17.6	0.5	3.4	--	0.8	5.0	*37.0*	1,469
NAIROBI at End Line														
Economic Level														
Lowest Level	62.2	55.6	4.2	10.7	2.3	26.0	7.6	1.0	3.0	1.8	0.9	6.6	*37.8*	256
Second Level	69.4	64.0	2.3	9.3	4.1	39.6	4.0	1.1	2.6	0.1	1.0	5.4	*30.6*	258
Middle Level	62.3	57.7	1.7	10.9	1.8	20.1	11.4	0.8	9.5	1.4	1.5	4.6	*37.7*	269
Second Highest	63.7	54.9	1.9	8.8	4.4	18.0	14.4	0.5	7.0	2.9	0.0	8.7	*36.3*	261
Highest Level	49.8	40.9	3.0	4.6	9.1	6.6	12.1	0.4	5.1	4.8	0.0	8.9	*50.2*	249
ALL LEVELS	61.6	54.8	2.6	8.9	4.3	22.1	9.9	0.7	5.5	2.2	0.7	6.8	*38.4*	1,294
SUBSET: WOMEN in a RELATNSHIP	75.8	68.3	3.5	12.7	6.2	29.5	13.2	0.2	2.3	2.6	0.6	7.5	*24.2*	787

Percentage Distribution with Respect to Economic Level and City Residence of Women (15 – 49 years of age) Currently using a Method of Contraception: MLE Project Baseline Study (2010) and End-line Study (2014)-- Mombasa, KENYA

	ANY METHOD	ANY MODERN METHOD	Sterili-zation	Implnts	IUD	Inject-ables	Pills	Emergncy Contra-ception	Male Con-doms	Stand. Days Methd	Other Modern Methd	ANY TRADITNAL METHOD	*NO METHOD*	N =
MOMBASA at Baseline														
Economic Level														
Lowest Level	27.8	23.2	0.8	1.4	0.4	13.1	4.3	0.0	2.4	--	0.8	4.6	*72.2*	256
Second Level	33.3	29.5	0.8	0.1	1.3	19.1	1.8	1.4	4.2	--	0.7	3.8	*66.7*	269
Middle Level	40.1	35.3	1.1	2.3	0.3	16.5	10.1	2.8	2.1	--	0.1	4.9	*59.9*	305
Second Highest	34.2	30.7	0.8	2.0	0.0	14.4	9.0	0.5	4.0	--	0.0	3.5	*65.8*	304
Highest Level	32.4	26.9	2.1	3.0	3.2	6.1	5.4	0.0	6.5	--	0.6	5.4	*67.6*	332
ALL LEVELS	33.7	29.3	1.2	1.8	1.1	13.6	6.3	0.9	3.9	--	0.4	4.5	*66.3*	1,465
SUBSET: WOMEN in a RELATNSHIP	48.4	41.5	1.4	2.6	1.3	20.6	10.6	1.3	3.0	--	0.7	6.9	*51.6*	837
MOMBASA at End Line														
Economic Level														
Lowest Level	43.7	42.5	1.1	11.8	0.2	17.0	3.3	0.1	7.6	0.0	1.2	1.2	*56.3*	160
Second Level	46.4	44.5	0.1	14.9	1.1	20.1	1.1	0.0	7.2	0.0	0.0	1.9	*53.6*	174
Middle Level	56.9	53.6	0.1	7.7	3.1	27.6	6.4	0.0	8.8	0.9	0.0	3.3	*43.1*	163
Second Highest	39.8	35.6	1.7	10.2	1.9	13.5	2.8	0.0	5.5	1.9	0.0	4.2	*60.2*	185
Highest Level	50.1	43.9	1.8	10.4	1.4	10.9	11.1	1.2	7.1	0.0	0.0	6.2	*49.9*	150
ALL LEVELS	47.1	43.8	0.9	11.0	1.6	17.8	4.7	0.2	7.2	0.6	0.2	3.3	*52.9*	832
SUBSET: WOMEN in a RELATNSHIP	52.8	49.7	0.9	14.0	2.1	22.3	6.2	0.3	3.4	0.3	0.4	3.1	*47.2*	558

Percentage Distribution with Respect to Economic Level and City Residence of Women (15 – 49 years of age) Currently using a Method of Contraception: MLE Project Baseline Study (2010) and End-line Study (2014) -- Kisumu, KENYA

	ANY METHOD	ANY MODERN METHOD	Sterili-zation	Implnts	IUD	Inject-ables	Pills	Emergncy Contra-ception	Male Con-doms	Stand. Days Methd	Other Modern Methd	ANY TRADITNAL METHOD	*NO METHOD*	N =
KISUMU at Baseline														
Economic Level														
Lowest Level	47.6	43.3	1.8	3.8	1.1	22.1	6.2	0.5	7.0	--	0.8	4.2	*52.4*	310
Second Level	50.0	48.4	1.8	8.1	0.0	26.2	4.5	0.5	6.9	--	0.4	1.6	*50.0*	302
Middle Level	47.9	44.5	1.2	2.7	0.8	22.9	5.5	0.3	10.2	--	0.8	3.5	*52.1*	298
Second Highest	44.5	41.0	1.7	3.7	0.6	20.7	5.0	0.6	8.3	--	0.4	3.5	*55.5*	327
Highest Level	49.6	45.1	3.2	4.7	2.9	13.6	4.2	1.4	14.6	--	0.5	4.5	*50.4*	366
ALL LEVELS	47.9	44.4	2.0	4.6	1.2	20.8	5.0	0.7	9.6	--	0.6	3.5	*52.1*	1,603
SUBSET: WOMEN in a RELATNSHIP	57.0	52.6	2.9	6.6	1.4	28.6	6.3	0.5	5.4	--	0.9	4.4	*43.0*	985
KISUMU at End Line														
Economic Level														
Lowest Level	61.5	58.3	3.1	29.8	0.8	19.4	0.9	0.0	3.2	0.0	1.2	3.2	*38.5*	189
Second Level	67.1	62.5	2.9	23.4	1.2	25.9	3.5	0.0	5.7	0.9	0.0	4.6	*32.9*	177
Middle Level	70.8	68.4	4.9	18.0	0.2	25.0	6.6	2.2	11.5	0.0	0.0	2.4	*29.2*	173
Second Highest	54.2	49.2	0.6	18.5	2.3	16.1	1.7	1.1	8.9	0.0	0.0	5.0	*45.8*	185
Highest Level	59.4	55.8	6.1	13.7	7.7	7.8	3.9	1.3	15.3	0.8	0.0	3.6	*40.6*	180
ALL LEVELS	62.5	58.7	3.5	20.8	2.4	18.8	3.3	0.9	8.9	0.3	0.2	3.8	*37.5*	904
SUBSET: WOMEN in a RELATNSHIP	73.1	68.4	4.0	25.4	3.2	24.3	4.2	0.4	6.7	0.5	0.2	4.7	*26.9*	593

Percentage Distribution with Respect to Economic Level and City Residence of Women (15 – 49 years of age) Currently using a Method of Contraception: MLE Project Baseline Study (2010) and End-line Study (2014) -- Machakos, KENYA

	ANY METHOD	ANY MODERN METHOD	Sterili-zation	Implnts	IUD	Inject-ables	Pills	Emergncy Contra-ception	Male Con-doms	Stand. Days Methd	Other Modern Methd	ANY TRADITNAL METHOD	*NO METHOD*	N =
MACHAKOS at Baseline														
Economic Level														
Lowest Level	52.7	44.3	6.3	3.4	0.5	23.8	4.5	0.4	5.2	--	0.2	8.4	*47.3*	283
Second Level	59.1	50.2	3.4	3.3	1.1	22.9	12.2	0.5	6.7	--	0.0	8.9	*40.9*	362
Middle Level	52.3	47.2	2.7	3.5	2.0	26.7	8.8	0.3	3.0	--	0.2	5.1	*47.7*	367
Second Highest	51.8	41.7	4.3	3.8	3.9	16.8	8.3	0.8	3.8	--	0.0	10.1	*48.2*	377
Highest Level	50.9	43.3	3.8	5.2	4.0	10.6	10.8	0.3	8.3	--	0.4	7.6	*49.1*	447
ALL LEVELS	53.2	45.3	4.0	3.9	2.5	19.5	9.2	0.5	5.5	--	0.2	8.0	*46.8*	1,834
SUBSET: WOMEN in a RELATNSHIP	72.9	63.6	6.6	6.0	3.6	28.8	15.3	0.2	2.7	--	0.3	9.3	*27.1*	997
MACHAKOS at End Line														
Economic Level														
Lowest Level	73.3	70.1	6.2	22.8	1.2	32.5	7.1	0.0	0.2	0.0	0.0	3.2	*26.7*	212
Second Level	63.7	58.0	4.1	10.0	2.9	24.8	10.1	1.1	4.8	0.5	0.0	5.7	*36.3*	221
Middle Level	59.7	55.0	1.8	9.9	1.8	30.3	8.5	0.9	1.6	0.0	0.2	4.7	*40.3*	232
Second Highest	66.4	56.5	3.2	8.1	5.7	21.4	13.2	0.4	3.8	0.0	0.8	9.9	*33.6*	235
Highest Level	60.6	50.3	4.5	8.7	9.8	9.0	13.5	1.3	3.1	0.9	0.3	10.3	*39.4*	226
ALL LEVELS	64.6	57.8	3.9	11.7	4.3	23.5	10.5	0.7	2.7	0.3	0.3	6.8	*35.4*	1,124
SUBSET: WOMEN in a RELATNSHIP	81.4	72.5	5.5	15.0	5.8	31.1	12.8	0.3	1.7	0.3	0.4	8.9	*18.6*	735

Percentage Distribution with Respect to Economic Level and City Residence of Women (15 – 49 years of age) Currently using a Method of Contraception: MLE Project Baseline Study (2010) and End-line Study (2014) – Kakamega, KENYA

	ANY METHOD	ANY MODERN METHOD	Sterili-zation	Implnts	IUD	Inject-ables	Pills	Emergncy Contra-ception	Male Con-doms	Stand. Days Methd	Other Modern Methd	ANY TRADITNAL METHOD	*NO METHOD*	N =
KAKAMEGA at Baseline														
Economic Level														
Lowest Level	46.2	44.6	1.1	1.8	1.9	29.6	2.4	0.0	6.1	--	1.7	1.6	*53.8*	232
Second Level	52.0	48.7	2.5	1.9	1.5	33.4	4.2	0.0	3.5	--	1.7	3.3	*48.0*	258
Middle Level	51.5	48.4	5.1	4.2	2.4	26.3	4.0	0.6	5.6	--	0.3	3.1	*48.5*	270
Second Highest	52.5	50.6	2.8	1.7	6.0	25.7	4.3	0.3	8.9	--	0.9	1.8	*47.5*	277
Highest Level	40.4	38.2	5.7	4.3	2.9	12.2	5.6	0.9	6.6	--	0.0	2.2	*59.6*	286
ALL LEVELS	48.5	46.1	3.6	2.8	3.0	25.1	4.2	0.4	6.2	--	0.9	2.4	*51.5*	1,324
SUBSET: WOMEN in a RELATNSHIP	61.2	57.9	4.8	4.1	3.9	34.1	5.4	0.0	4.6	--	1.1	3.3	*38.8*	826
KAKAMEGA at End Line														
Economic Level														
Lowest Level	59.4	58.3	2.4	26.2	1.8	21.5	3.3	0.0	3.1	0.0	0.0	1.1	*40.6*	167
Second Level	66.2	65.5	3.8	11.6	1.3	38.7	5.6	0.0	3.1	0.0	1.3	0.7	*33.8*	147
Middle Level	55.1	54.8	2.6	15.1	2.2	25.8	4.6	0.0	3.1	0.0	1.6	0.3	*44.9*	172
Second Highest	48.4	47.5	3.2	9.5	5.0	15.1	6.9	0.0	7.4	0.0	0.3	0.9	*51.6*	180
Highest Level	49.8	44.6	6.4	11.0	5.8	7.4	8.4	0.4	5.2	0.0	0.0	5.2	*50.2*	160
ALL LEVELS	55.5	53.8	3.6	14.7	3.3	21.3	5.7	0.1	4.4	0.0	0.6	1.6	*44.5*	826
SUBSET: WOMEN in a RELATNSHIP	64.7	62.4	4.5	17.9	4.1	25.8	7.2	0.1	1.9	0.0	0.8	2.3	*35.3*	571

Thank You!

Lunch Break (60 minutes)

Training in Data Visualization

Day 2, Afternoon –

MLE

Process Step:
Chapter 3

Developing a Data Visualization

- **Learning Activity** – Presentation on the other types of graphical data visualizations (beyond those found in a written report) and the considerations of design for each type:
  - Infographics
  - Conference Poster
  - Dashboard
  - Powerpoint Presentation
  - Web-based video
  - Film

(30 minutes)

MLE

Developing a Data Visualization:
Different Types of Data Visualizations

As objectives, at the end of this session, participants will be able to:

- Identify the different types of graphical data visualizations
- Identify the different design considerations in the development of each type of graphical data visualization
- Evaluate several different types of data visualizations

(30 minutes)

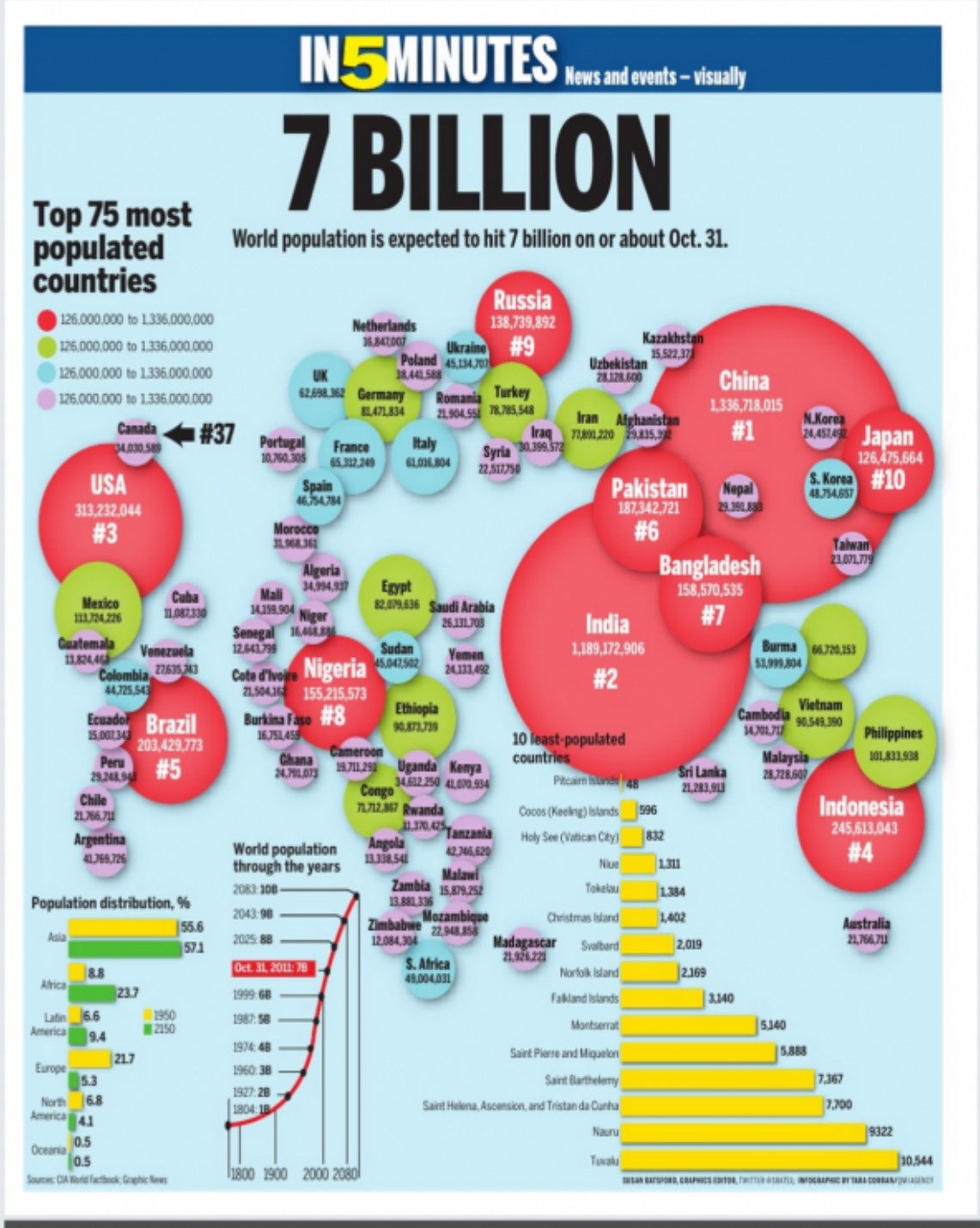
MLE

Developing a Data Visualization:
Types of Graphical Data Visualizations and Design Considerations for Each Type (1)

- **Infographics**
  - **Logical flow** in the presentation of several graphical elements and text
  - **Effective/harmonious relationship** among these elements and the text regarding **size** and **positioning**
  - **Meaningful use of colors** and **font** characteristics (**size** and **style**)

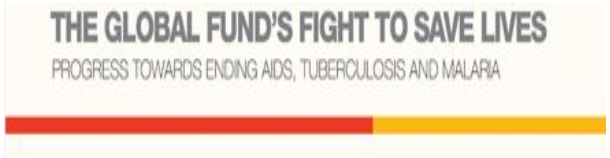
MLE

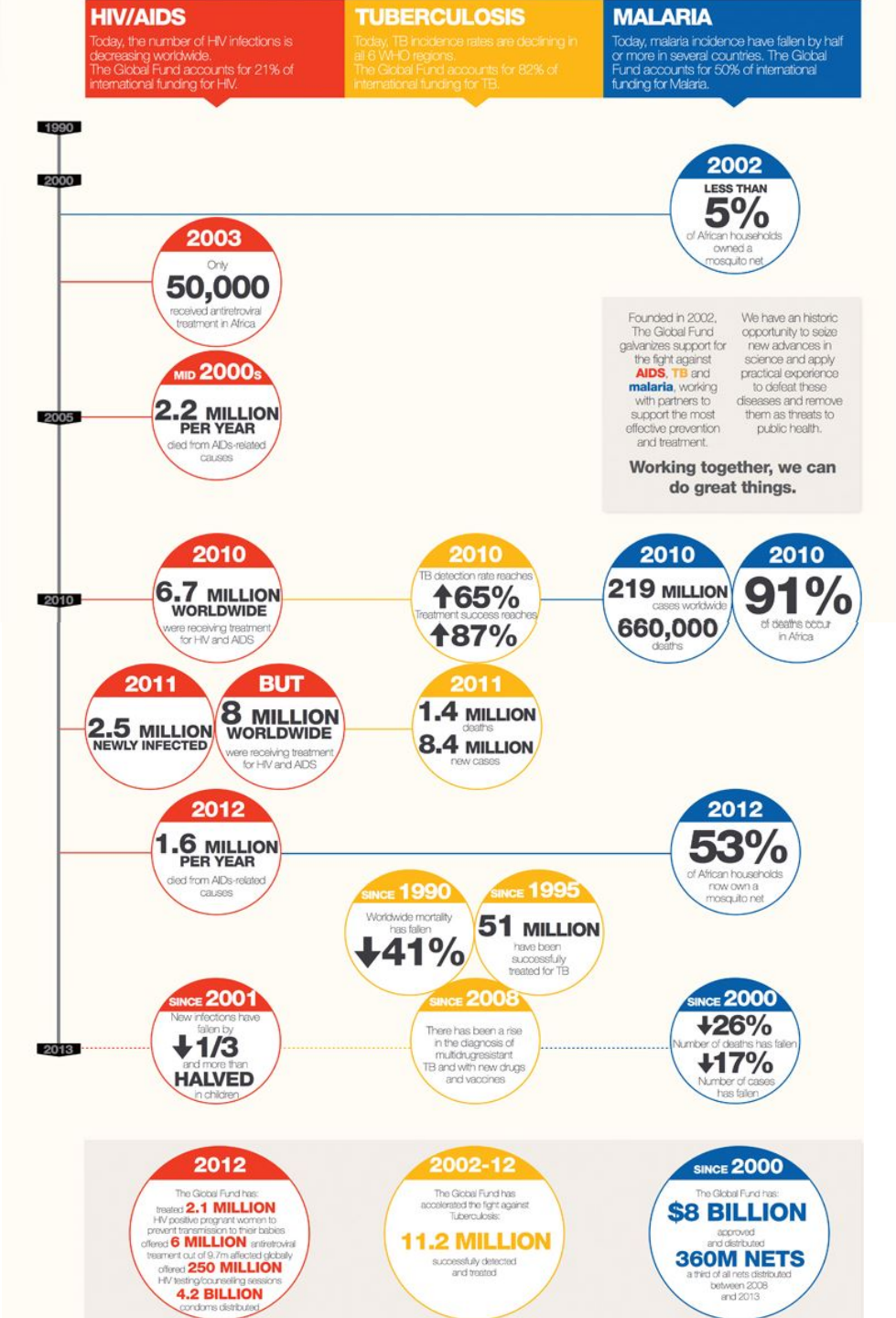
- Example of an Infographic



MLE

- Another Example of an Infographic





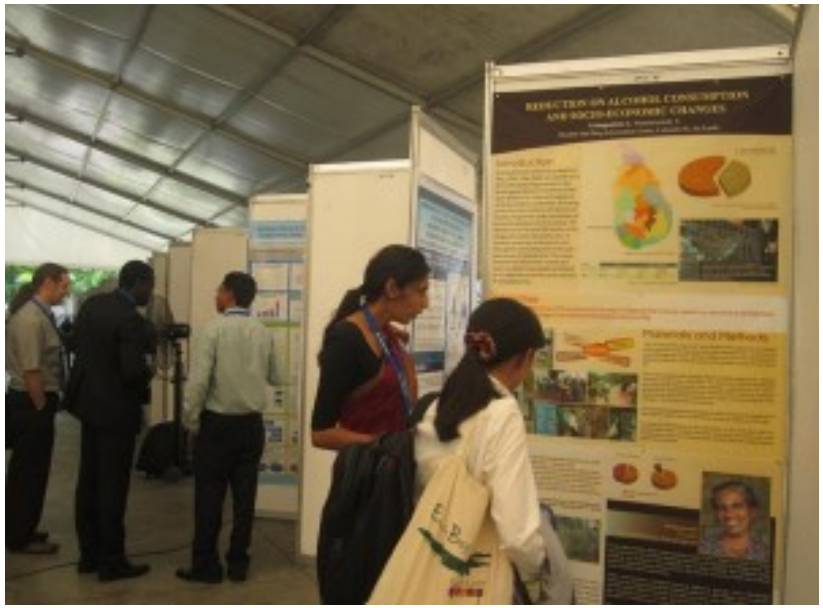
MLE

Developing a Data Visualization:
Types of Graphical Data Visualizations and Design Considerations for Each Type (2)

- **Conference Poster**
  - **Same considerations** as those for infographics:
    - logical flow in the presentation of several graphical elements and text
    - effective/harmonious relationship among these elements and the text regarding size and positioning
    - meaningful use of colors and font characteristics (size and style)
  - **Scale of size** of the poster itself

MLE

- Examples of Conference Posters



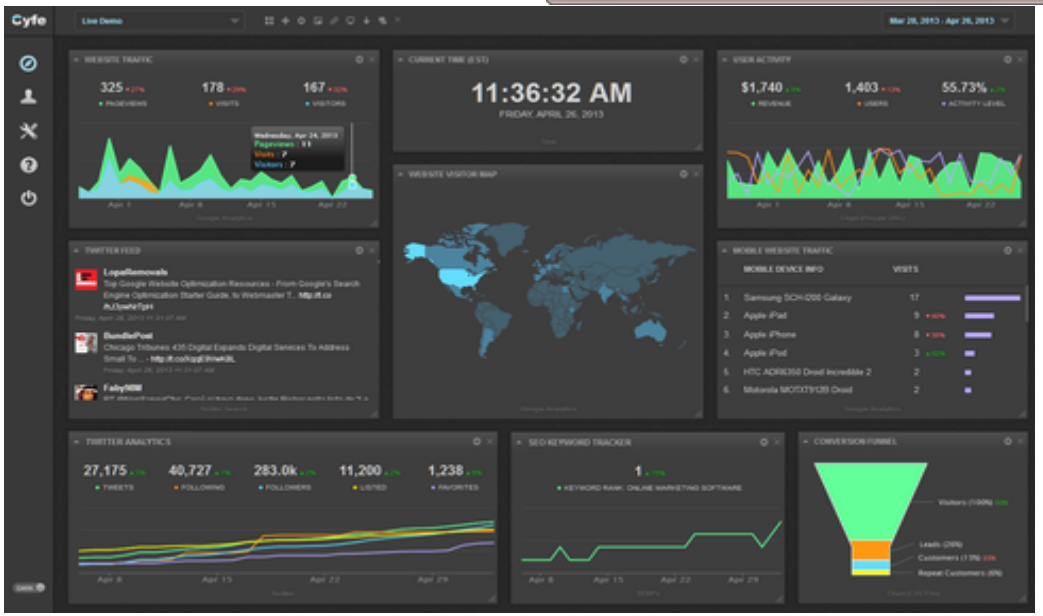
MLE

Developing a Data Visualization:
Types of Graphical Data Visualizations and Design Considerations for Each Type (3)

- **Dashboard**
  - Same considerations as those for infographics and conference posters:
    - logical flow in the presentation of several graphical elements and text
    - effective/harmonious relationship among these elements and the text regarding size and positioning
    - meaningful use of colors and font characteristics (size and style)
    - scale of size of the computer screen
  - **Flexibility in data sorting and filtering**

MLE

- Examples of Dashboards



Developing a Data Visualization:
Types of Graphical Data Visualizations and Design Considerations for Each Type (4)

- **PowerPoint Presentations**
  - Same considerations as those for infographics, conference posters, and dashboards:
    - logical flow in the presentation of several graphical elements and text
    - effective/harmonious relationship among these elements and the text regarding size and positioning
    - meaningful use of colors and font characteristics (size and style)
    - scale of size of the projector screen
  - **Distance of spectators** from the **projector screen**

MLE

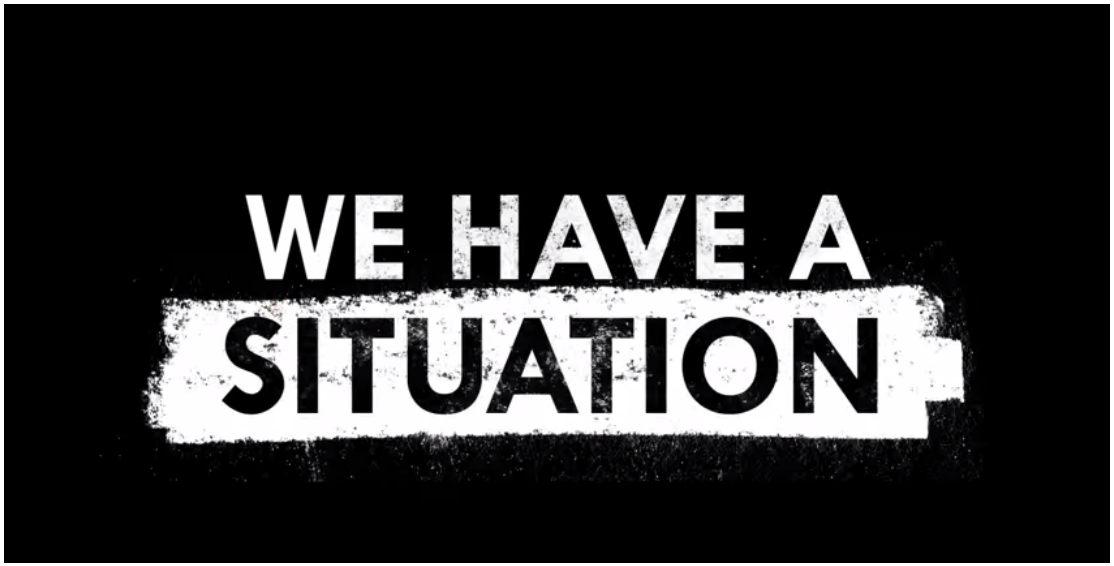
Developing a Data Visualization:
Types of Graphical Data Visualizations and Design Considerations for Each Type (5)

- **Web-based Video**
  - Same considerations as those for infographics, conference posters, dashboards, and PowerPoint:
    - logical flow and **good pacing** in the presentation of several graphical elements and text
    - effective/harmonious relationship among these elements and the text regarding size and positioning
    - meaningful use of colors and font characteristics (size and style) and, as well, **use of music and other sounds**
    - scale of size of the computer screen

MLE

- Examples of web-based videos

https://www.youtube.com/watch?v=Uyu7vGguWBQ



https://www.youtube.com/watch?v=1e8xgF0JtVg

MLE

Developing a Data Visualization:
Types of Graphical Data Visualizations and Design Considerations for Each Type (6)

- **Film**
  - Same considerations as those for infographics, conference posters, dashboards, PowerPoint presentations, and web-based videos:
    - logical flow and good pacing in the presentation of several graphical elements and text
    - effective/harmonious relationship among these elements and the text regarding size and positioning
    - meaningful use of colors and font characteristics (size and style) and, as well, use of music and other sounds
    - **scale of size** of **movie screen**

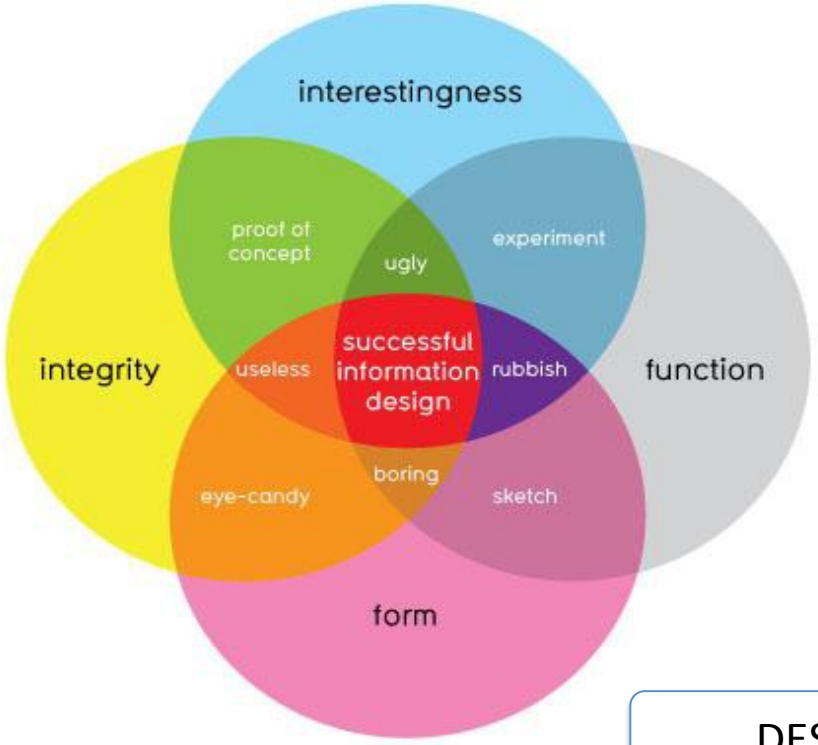
MLE

Developing a Data Visualization:

Four Essential Needs for a Good Visualization: **Function, Form, Integrity, and Visual Appeal**

INFORMATION

relevance
meaningfulness
novelty



truth
consistency
honesty
accuracy

ease of use
usefulness
usability
fit

DESIGN

beauty
structure
appearance

MLE

Coffee / Tea Break (30 minutes)

Process: Developing a Data Visualization

1. Identify your audience & context
- Who are the stakeholders?
- What do they need to know?

2. Find the story in your data
- What information do you have?
- Are there identifiable patterns, trends, surprises, relationships, successes or failures?

3. Build your Visualization
- Sketch first!
- Choose the right type of chart for your data
- Simplify
- Use pre-attentive attributes to highlight your story

4. Disseminate, share, and use!
- How will this visualization be used?
- What is the ideal format?
- Good data visualization ensures that data are read, understood, and used for evidence-based decision-making.

MLE

Process:
Step 4

Disseminating a Data Visualization and Assessing the Results

- **Learning Activity 1** – Brief review of good practice measures before full dissemination of a data visualization to the target audience
  (40 minutes)

MLE

Considerations (1)

PURPOSE

- As described in early part of this presentation, a data visualization is designed to transform key findings of data into a key message for a selected target audience through a visual presentation that:

  ✓ is easily intelligible and persuasive to that audience

  ✓ fosters a compelling response to the key message by that audience

MLE

Considerations (2)

PERCEPTION

- A data visualization should be tested with a sample set of target audience members for their level of comprehension and response to it — before dissemination to the full target audience — to determine
  - ✓ if the key message provided by the data visualization is well understood by the audience
  - ✓ whether or not the expected response to the key message was elicited in the audience

MLE

Considerations (3)

PLAN for Dissemination

Purpose

- ✓ what are the major findings your data visualization is expected to show to your target audience?
  - how can the level of success in communicating these findings be measured?
- ✓ what are the expected responses by your target audience after « reading » this data visualization?
  - how can the level of success in eliciting these responses be measured?
  - how can the level of responsive action be measured?

MLE

Considerations (4)

PLAN for Dissemination

People

- ✓ what are measurable characteristics of your target audience?
  - how many of the total number of target audience members are expected to be reached by the data visualization?
  - how suitably diverse/representative should those audience members be who are reached?
  - how reliable are levels of numeracy and technical skills that are assumed to be held by target audience members by which to read the data visualization?

MLE

Considerations (5)

PLAN for Dissemination

Process

- ✓ was there ongoing self-assessment during the steps to develop a data visualization?
  - how much diversity in skills and technical backgrounds is there among those involved in developing the data visualization?
  - how complete and accurate is the source data and its analysis for the data visualization?
  - how rigorous was the collective review of data to identify key findings that would be especially relevant to the target audience
  - how thoroughly self-critical were the developers of the data visualization to ensure that the « message » to be « read » was made clearly, concisely, and memorably?

MLE

Considerations (6)

If the test results of the data visualization are found to be **unsatisfactory**, then review assumptions and decisions made in:

- ✓ identifying and describing the appropriate target audience and the expected response by its members to the data visualization (**People**)
- ✓ selecting the type, content, and design of the data visualization (**Process**)

- ○ Be prepared to revise earlier assumptions and decisions

MLE

Considerations (7)

PEOPLE

- Is there a mismatch between the target audience selected and the content and design of this data visualization?
- Were target audience members correctly assessed on their:
  - ✓ knowledge of topics presented by the data visualization (generalists or specialists)?
  - ✓ numeracy and related technical skills to « read » the data visualization?

MLE

Considerations (8)

PEOPLE

- Were target audience members correctly assessed on their:
  - ✓ interest in learning about the key findings?
  - ✓ willingness to invest time and attention to learn about them?
  - ✓ reasons for wanting to learn about the key findings?
  - ✓ level of responsiveness to the key findings?
  - ✓ level of access to the data visualization?

MLE

Considerations (9)

PEOPLE

- Was there an appropriate mix of individuals involved in the data visualization design, with storyboard participation by:
  - ✓ senior and junior technical staff
  - ✓ knowledge management staff
  - ✓ graphic design staff
  - ✓ field staff who work directly with members of the target audience

MLE

Considerations (10)

PROCESSES

- In preparation for the data visualization, has there been:
  - ✓ thorough check of the quality of the data?
  - ✓ careful selection of context for the interpretation of the data — pattern, trend, comparison, range, outliers, surprises?
  - ✓ rigorous search for the « data story » through the use of the storyboard method — group discussions, sketched ideas, synthesis of data results into key findings?

MLE

Considerations (11)

PROCESSES

- In preparation for the data visualization, has there been:
  - ✓ selection of the most appropriate graphics/charts to highlight the key findings?
  - ✓ success in designing the graphics/charts with simplicity, honesty and with key use of pre-attentive attributes?

MLE

Reflections

- Discuss as a group a summary of the topics covered during the course of today (20 minutes)

- Separate yourselves into pairs in order to evaluate the quality of instruction of the morning's and afternoon's session by means of the graphic method of assessment that was developed by each pair
  (10 minutes)

MLE

Reflections

- What else remains to be learned about data visualization?
  - Topics for future study

Global Health eLearning Center

Data Visualization - An Introduction

https://www.globalhealthlearning.org/course/data-visualization-brave-new-world

BetterEvaluation

http://betterevaluation.org/plan/describe/visualise_data

CONNECT

DataVizHub.co

MLE

Gurus of Data Visualization:

Dr. Edward Tufte

"The Visual Display of Quantitative Information"; "Visual Explanations"; "Beautiful Evidence"

https://www.youtube.com/watch?v=Th_1azZA2OY "Beautiful Evidence"

Hans Rosling (YouTube, TED Talks)

"200 Countries, 200 years , 4 Minutes": https://www.youtube.com/watch?v=jbkSRLYSojo

"Joy of Stats": https://vimeo.com/18477762

Steven Few (business analytics, dashboards)

"Now You See It: Simple Visualization Techniques for Quantitative Analysis"

Andrew Kirk (journalist)

"8 Hats of Data Visualization Design"

https://vimeo.com/44886980

Alberto Cairo (journalist; MOOC on data visualization)

"The Functional Art"

http://www.thefunctionalart.com/

Cole Nussbaumer

Storytelling with Data

MLE

Stephanie Evergreen

Consulting Site: http://stephanieevergreen.com/

"Presenting Data Effectively: Communicating Your Findings for Maximum Impact" (Sage Publisher, October 2013)

Ann K. Emery

Blog: http://annkemery.com/

http://annkemery.com/essentials/

Jon Schwabish (Urban Institute, Congressional Budget Office)

Blog: http://blog.visual.ly/author/jon-schwabish/

http://policyviz.com/presentations/graphic-continuum-video/

Other Notables in Data Visualization:

8 Great Books about Data Visualisation (Andy Cotgreave)"

http://www.tableau.com/about/blog/2013/7/list-books-about-data-visualisation-24182

David Giard

Effective Data Visualization: https://www.youtube.com/watch?v=nP6qWhOkha4

Nancy Duarte

Diagrammer (4,000 customizable diagrams to download for free with which to enhance your PowerPoint presentations): http://www.duarte.com/diagrammer/

MLE

Thank You!

MLE