Communicating Vital Statistics Through Visualizations

Workshop on Vital Statistics for North and Central Asian Countries
Bishkek, Kyrgyzstan, 7-11 October 2019
Session objectives

By the end of the session, participants will be able to:

• Describe and compare the main types of data visualization

• Identify the factors involved in choosing the type of data visualization

• List design principles that contribute to effective data visualization

• Visualize vital statistics data using maps
Factors in Choosing Visualization Type

• Communication Purpose
  • Change
  • Comparison
  • Composition
  • Correlation

• Characteristics of Data
  • Number of series displayed
  • Number of points displayed within each series
Line Graph
Matching Visualization to Purpose and Data

Communication Purpose:
*I want to show the change over time in life expectancy*

Characteristics of Data:
*I want to show one series with many data points*
Matching Visualization to Purpose and Data

Communication Purpose:
*I want to show the change over time in life expectancy*
*I also want to compare values across sex*

Characteristics of Data:
*I want to show two series with many data points*
Visualization Type: Line

Figure B18.4 Life expectancy at birth for males and females in Norway, 1850-2015

Source: Statistics Norway statistics bank.
Visualization Type: Line

Figure 7: Age Specific Mortality Rates by period, (2015-2017)

Matching Visualization to Purpose and Data

Communication Purpose:

*I want to compare values for mortality rates across categories*

Characteristics of Data:

*I want to show rates for five groups (five series)*

*I want to show information for two years (two data points for each series)*
Figure 13. Age-adjusted Premature Death (Age <65 years) Rates by Neighborhood Poverty*, New York City Residents, 2008 and 2017

Source: New York City Department of Health and Mental Hygiene
Matching Visualization to Purpose and Data

Communication Purpose:
I want to compare values for total fertility rates across regions

Characteristics of Data:
I want to show rates for only one group (one series)
I want to show information for 15 regions (15 data points)
Visualization Type: Bar

Figure 4.2: Crude Birth Rate by Province, Zambia 2016

- Urban: 8.0
- Rural: 4.8
- Zambia: 6.2
- Central: 11.7
- Lusaka: 11.3
- Luapula: 7.8
- Copperbelt: 6.1
- Eastern: 4.0
- Southern: 3.7
- Western: 3.4
- Muchinga: 2.3
- North Western: 1.8
- Northern: 1.6

Source: Department of National Registration Passport and Citizenship (DNRPC) 2013-16 Administrative Data
Stacked Column/Bar
Communication Purpose:
I want to break down causes of death (composition)

Characteristics of Data:
I want to show rates for two time period (two series)
I want to show three broad groups of causes (three data points for each series)
Visualization Type: Stacked Column

CAUSES OF DEATH
WORLD-WIDE, 2000 AND 2016

Source: World Health Organization
CAUSES OF DEATH WORLD-WIDE, 2000 AND 2016

- **Injuries**: 45.5 (2000), 48.8 (2016)
- **Non-communicable diseases**: 316.3 (2000), 405.5 (2016)
- **Communicable diseases**: 161.2 (2000), 114.5 (2016)

Source: World Health Organization
Scatter
Matching Visualization to Purpose and Data

Communication Purpose:
I want to show the correlation between TB death rates and HIV death rates.

Characteristics of Data:
I want to show rates for ten different countries (ten data points).
HIV and TB Death Rates in Eastern Europe, 2012

Source: World Health Organization mortality database
## Design Principles

<table>
<thead>
<tr>
<th>Guide Viewer</th>
<th>Eliminate Distractions</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Label sufficiently</td>
<td>- Present text as it will be scanned</td>
</tr>
<tr>
<td>- Visually link related elements</td>
<td>- Limit non-data elements</td>
</tr>
<tr>
<td>- Create a visual hierarchy</td>
<td>- Use formatting purposively</td>
</tr>
<tr>
<td>- Simplify data comparisons</td>
<td>- Be cautious with images</td>
</tr>
</tbody>
</table>
Mapping Vital Statistics
Importance of geography — why map?

- Relating data to location as powerful analysis
- Visualizing health outcomes by geography
- Identifying geographic trends


Figure 11.2: Distribution of Hospital Live Births by place of occurrence in Sri Lanka, 2016
Source: Medical Statistics unit
Choropleth Maps

• Used for **prevalence, standardized rates** and **ratios** linked to administrative areas

• Division of data into **categories**
  • Rankings from *high to low* or *low to high*
  • Number of categories from 3–6
Crude Deaths Rates – Mumbai, 2015

Legends

• A legend defines symbols and/or colors important to the map
  • Information necessary for reader

• Not all legend pieces are needed in map
  • Bar scales necessary if distance is important
  • If map does not point true north, a compass can be added for orientation if important
Crude Deaths Rates – Mumbai, 2015

What is Needed for Mapping

• Data for geographical area
  • Shapefiles for areas to be mapped
    • Administrative areas for choropleth maps

• Health data or events linked to location
  • Latitude/longitude of events
  • General location
  • Addresses for geocoding

• Software
Limitations of Mapping

• Reliance on spatial data
• Cannot show all factors relevant to health issue
• Cannot convey all information necessary for understanding health issue
Summary

• When choosing and creating visualizations, consider:
  • The story you want to tell
  • Communication purpose
  • Characteristics of data
  • Design principles

• Mapping health information can be a compelling visual method
Acknowledgements

• Bloomberg Philanthropies Data for Health Initiative
• Vital Strategies
• US Centers for Disease Control and Prevention
• University of Melbourne
• Statistics Norway
• ESCAP
• EFTA