visualizing health and healthcare data

Creating clear and compelling visualizations to "see how you're doing"

katherine **rowell**
lindsay **betzendahl**
cambria **brown**

WILEY

Visualizing Health and Healthcare Data

Visualizing Health and Healthcare Data

Creating Clear and Compelling Visualizations to "See How You're Doing"

Katherine Rowell
Lindsay Betzendahl
Cambria Brown

WILEY

Copyright © 2021 by HealthDataViz. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600, or on the Web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993, or fax (317) 572-4002.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at http://booksupport.wiley.com. For more information about Wiley products, visit www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Names: Rowell, Katherine L., author. | Betzendahl, Lindsay, author. | Brown, Cambria, author.

Title: Visualizing health and healthcare data : creating clear and compelling visualizations to "see how you're doing" / by Katherine Rowell, Lindsay Betzendahl, Cambria Brown.

Description: First edition. | Hoboken, New Jersey : Wiley, [2020] | Includes bibliographical references and index.

Identifiers: LCCN 2020032174 (print) | LCCN 2020032175 (ebook) | ISBN 9781119680888 (paperback) | ISBN 9781119682325 (adobe pdf) | ISBN 9781119680864 (epub)

Subjects: LCSH: Medical statistics—Data processing. | Information visualization.

Classification: LCC RA409.5 .R69 2020 (print) | LCC RA409.5 (ebook) | DDC 610.2/1—dc23

LC record available at https://lccn.loc.gov/2020032174

LC ebook record available at https://lccn.loc.gov/2020032175

Cover Design and artwork: Wiley

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

Contents

	Preface	xiii
Section I	**Establishing a Framework and Process**	**1**
Chapter 1	Health and Healthcare Data Visualizations of Historical Importance	3
Chapter 2	Stop Hunting Unicorns and Start Building Teams and Know The Data	9
	Search for Characteristics and Core Competencies	10
	Get to Know the Data	11
	Classifications, Intent, Purpose, and Lineage	12
	Two Types of Data	14
	Qualitative/Categorical Data	14
	Quantitative/Numerical Data	14
	Scales/Levels of Measure	15
	Nominal	15
	Ordinal	16
	Interval	17
	Ratio	18
	Summary	19

Chapter 3 Requirements-Gathering and Design Methods 21

Design Thinking Foundational Concepts 22
Design Methods 23
Contextual Inquiry 23
Mental Models 24
Personas 26
Persona Creation Guide 27
Graphic Organizers 29
Guided Analytics Framework 29
Summary Overview Dashboard 30
Supporting Focused Reports 30
Details 31
Multidimensional Exploratory Displays (MEDs™) 31
Sketching 32
Prototyping 33
Testing 34
Summary 36

Section II Perceiving the Best Practices of Data Visualization 37

Chapter 4 The Research 39

Research Informs Data Visualization Best Practices 39
Preattentive Attributes 41
Preattentive Attributes at Work 43
Gestalt Principles 46
Color Theory 48
The Power of White Space 53
Where People Look 54
Summary 54

Chapter 5 Table Design Checklist 55

Fundamentals of Table Design 55
Organization/Categorization 55
Non-Data Ink 56

Fonts 56
Number Alignment and Formatting 56
Labels 58
Summary 58

Chapter 6 Powerful Visualizations in Four Shapes 59
Bars, Lines, Points, and Boxes 59
Shape One: Bars 61
Bar Basics 61
Using Bars To: See How You're Doing 63
Distributions 63
Histograms 63
Population Pyramid (Paired Bars) 65
Ranking 67
Change over Time 68
Comparing Multiple Data Points 70
Proportions | Part-to-Whole 71
Challenging the 100% Myth 73
Deviation (Difference, Variation) 73
Ranges and Comparative Values 74
Displaying the Vital Few: Pareto Charts 77
Bars Are Not Boring 78
Shape Two: Lines 79
Line Basics 79
Using Lines To: See How You're Doing 83
As a Reference | Comparison 83
Change over Time 84
Change over Time | Sparklines 86
Change over Time | Deviation Graphs 87
Distributions 88
Distributions | The Empirical Rule and Control Charts 88
Statistical Process Control Charts (SPCs) and Geometric (G) Charts 90
Relationships | Correlations 91
Shape Three: Points 92
Point Basics 92

Using Points To: See How You're Doing 94
Distributions 94
Revealing Details 95
Change over Time 96
Correlation 96
Hierarchy Quadrant 97
Location Details 98
Shape Four: Boxes 98
Box Basics 99
Using Boxes To: See How You're Doing 99
Distribution 99
Multiple Values 100
Change over Time and Utilization Rates 101
Hierarchical Data 101
Other Shapes 102
Summary 103

Chapter 7 Maps 105

Using Maps to Gain Insights 105
Geographic Maps 105
Choropleth Maps 106
Hex-Tile Maps 109
Symbol/Dot-Density Maps 110
Proportional Symbol Maps 112
When Not to Use a Map 113
Summary 114

Chapter 8 Graphs and Charts to Never Use or Use with Caution 115

When "Cool Displays" Are Anything But 115
Pie and Donut Charts 117
Why People Use Them 117
Characteristics 118
Challenges 118
Best Practice Alternative 119

Multiples of Several-Part Stacked Bar Charts (MSPSBCs) 121
Why People Use Them 121
Characteristics 121
Challenges 121
Best Practice Alternative 124
Bubble Charts 124
Why People Use Them 124
Characteristics 125
Challenges 125
Best Practice Alternative 126
Treemaps 128
Why People Use Them 128
Characteristics 129
Challenges 129
Best Practice Alternative 130
Marimekko (Mekko or Mosaic) Charts 132
Why People Use Them 132
Characteristics 132
Challenges 134
Best Practice Alternative 134
Radial Bar and Petal Charts 134
Why People Use Them 134
Characteristics 135
Challenges 135
Best Practice Alternative 136
Radar Charts 138
Why People Use Them 138
Characteristics 138
Challenges 138
Best Practice Alternative 139
Sankey Diagrams 141
Why People Use Them 141
Characteristics 141
Challenges 142
Best Practice Alternative 144

One More Thing: 3-D 145
Summary 146

Chapter 9 Making Accessible Visualizations 149
Accessible Design Is Good Design 149
Accessibility in Data Visualization 150
Ways to Make Accessible Data Visualizations 151
Summary 158

Section III Creating Compelling Data Displays 159

Chapter 10 Dashboards, Reports, and Multidimensional Exploratory Displays (MEDs™) 161
Definitions Matter 161
Dashboards 162
Dashboards Defined 162
Purpose/Objective 162
Data/Information 163
Design 164
Example Dashboards 165
Dashboard Summary 169
Reports 170
Reports Defined 170
Purpose/Objective 170
Design 170
Example Reports 171
Report Summary 176
Multidimensional Exploratory Displays (MEDs™) 176
MEDs™ Defined 177
Purpose/Objective 177
Design 177
Example MED™ 177
MEDs™ Summary 184
Summary 184

Chapter 11 Infographics 185
"No Tobacco Day" Infographic 186
Measles and Vaccinations Infographic 188
Infographic vs. Infoposter 191
Summary 194

Section IV Closing Thoughts and Recommended Reading and Resources 195

Closing Thoughts 197
Fluency and Mastery 197

Bitten by the Viz Bug | Recommended Reading and Resources 199
Recommended Reading 199
Resources 201
Accessibility Resources 202

Author Bios 203

References 207

Index 211

Preface

We understand the challenges you face when trying to communicate health and healthcare data. We wrote this book to share our deep experience and expertise in creating clear, compelling, and actionable displays of health and healthcare data that empower people to "see how they're doing."

As a group, we've spent the last 30 years working on this book—29 conducting the research, and one writing it. Our research can be categorized into three areas: formal education, work experience, and self-education.

We've each had the privilege of a formal education delivered by recognized health and healthcare experts and thought leaders. This education is foundational to our ability to think critically about the various aspects of what health is, how health and healthcare systems operate, and the structures and policies that influence it all.

We've played a role in the development, capture, and analysis of just about every type of health and healthcare data imaginable—from administrative claims data to risk-adjusted clinical outcomes data to complex public health survey data. We've deciphered the inner workings of innumerable transactional systems and untangled databases created by evil geniuses.

Each of us can describe the moment we were bitten by the data visualization bug and set on a path of self-education. We sought out thought leaders in the field—some we know only through their books and blogs; others we have developed lifelong professional and personal friendships with. We learned and honed our viz skills through deep practice—stretching ourselves outside our comfort zone, stopping to reflect on successes and errors, making adjustments, and continuing the process over time. As a result, we have proudly joined the ranks of recognized data visualization experts, and we work with leading health and healthcare organizations throughout the world. And now, we have written this book to share our experience and expertise with you.

This book is organized into three sections. Section I is focused on understanding different types of data to be displayed, and requirement-gathering methods. All too often teams want to jump directly to creating displays, but this is a mistake to be avoided. Before you can determine how to display data and information, you must understand what will be displayed and for whom.

Section II provides an overview of the research that informs the best practices of table and graph design. Included in this chapter are the four fundamental shapes you can use alone or in combination to create clear and compelling displays of data. This section also includes detailed examples of the most common mistakes people make in selecting a chart type, with explanations about why they don't work, and examples of what works better and why.

Section III of the book defines and describes the characteristics of dashboards, reports, multidimensional exploratory displays (MEDS™) and infographics. It includes examples of each type of display and tips and tricks to designing and building them.

We ardently believe that making the message and the story in the data clear will improve health and healthcare systems. We hope this book will help you to join in our efforts.

SECTION ONE

Establishing a Framework and Process

"If you can't describe what you're doing as a process, you don't know what you're doing."

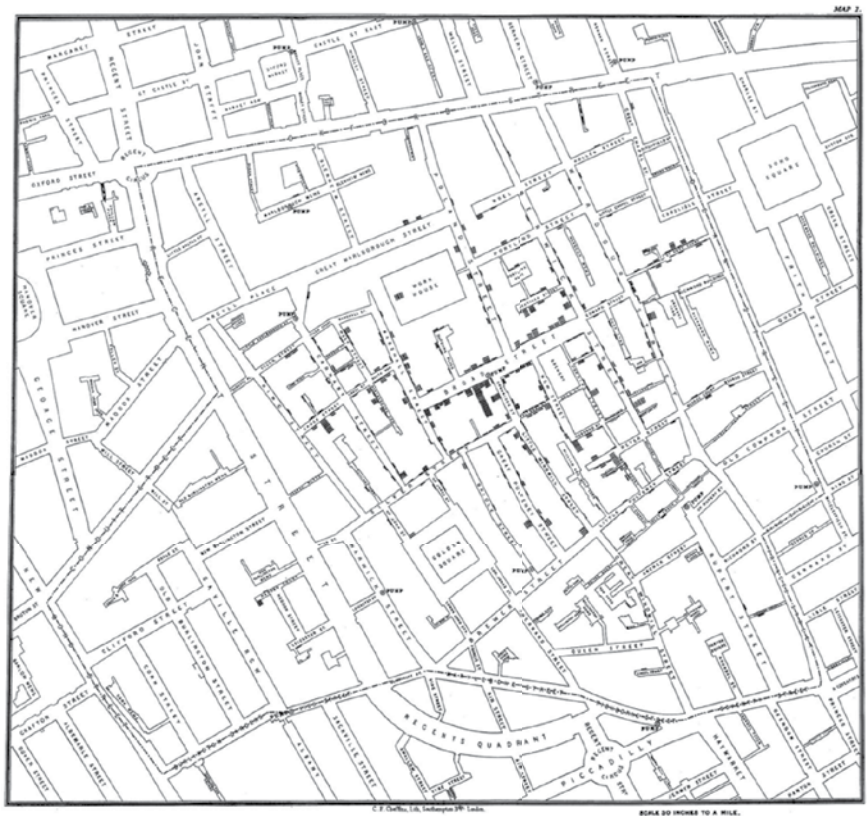
—W. Edwards Deming,
American engineer,
statistician, and leading quality
management thinker

Chapter One

Health and Healthcare Data Visualizations of Historical Importance

Even before modern-day visualization research validated the direct and powerful relationship between the way information is presented and the way we see and understand it, pioneering healthcare statisticians and caregivers like John Snow (1813–1858) and Florence Nightingale (1820–1910) understood that visual display could be a highly effective method for grasping and communicating the messages buried in data. No one who has ever taken an epidemiology course can forget Dr. John Snow's classic work, *On the Mode of Communication of Cholera*. By mapping the London street addresses of residents who had become sick (and in many cases died) and their distance from City water pumps, Snow could visually and effectively communicate the relationship between a single pathogen-tainted water source and the homes of people who contracted the disease. Most people who had fallen ill, it turned

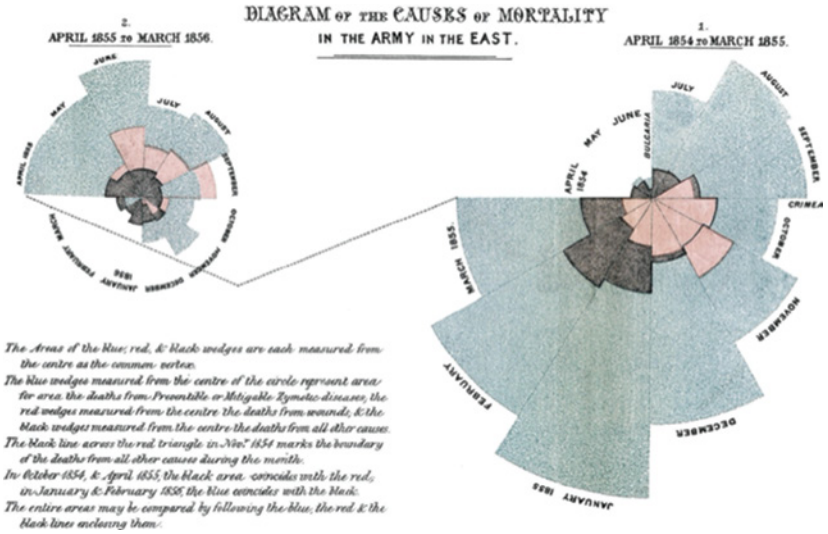
out, lived near the Broad Street pump. Snow persuaded the town council to remove the pump's handle, and the outbreak abated.



In 1868, British nurse Florence Nightingale—distressed by the alarmingly high mortality rates in the Crimean War—began to compile statistics on causes of death. Her analysis revealed that of the 900,000 soldiers who died during the war—more than half of 1,650,000 combatants from all countries involved—most had succumbed to preventable diseases arising from unsanitary conditions in the hospitals where they were treated, and not as a direct result of battlefield wounds. Nightingale recognized the buried message: better hygiene could have saved—and could still save—thousands of lives.

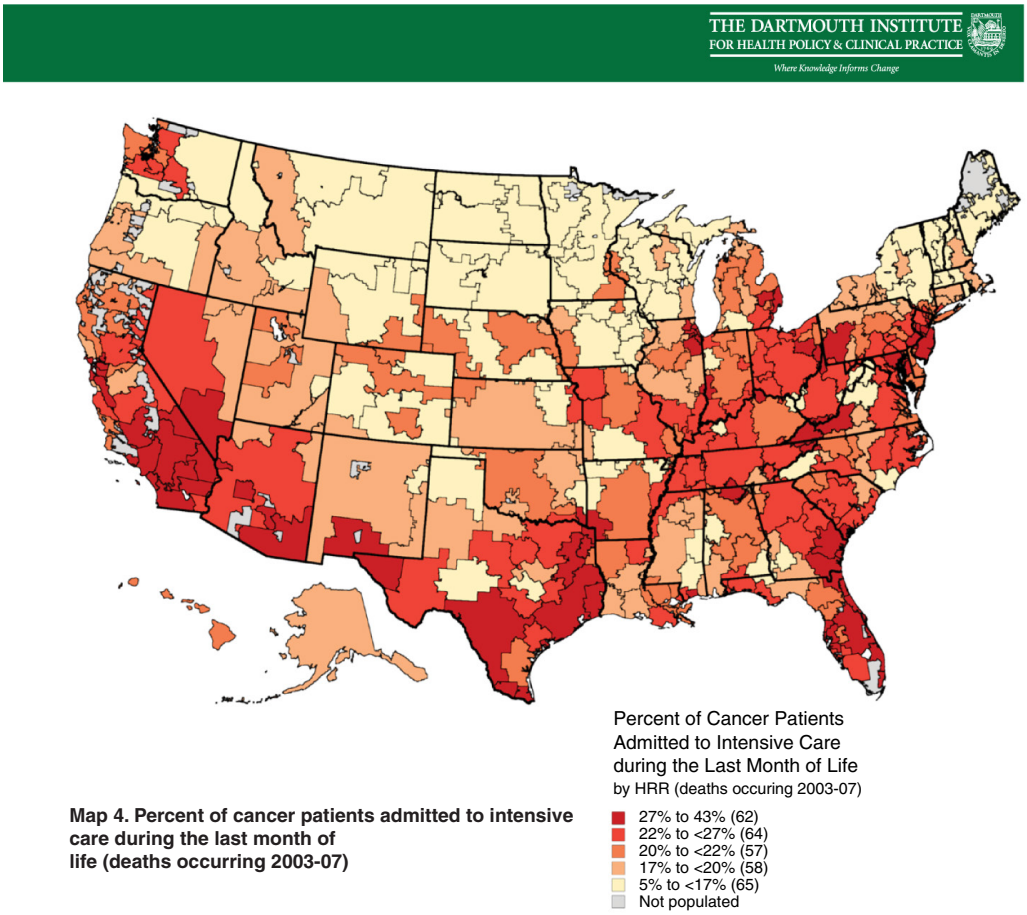
As impressive as her statistics were, Nightingale worried that the tables she presented to Queen Victoria would seem tedious,

even incomprehensible, and feared that members of the British Parliament were unlikely to be swayed by numbers lying flat on a page. So Nightingale devised ingenious ways of presenting the information in charts.



In the now-famous "Diagram of the Causes of Mortality in the Army in the East," each month is represented as a twelfth of a circle. The months with more deaths are shown with longer wedges so that the area of each wedge represents the total number of deaths. Preventable deaths are blue, deaths due to wounds are red, and deaths from all other causes are black. Over the months after March 1855, when members of the Sanitary Commission began repairing, cleaning, and otherwise improving field hospital conditions, the blue wedges shrank dramatically. Showing incredible insight into the power of displaying the data in this way, Nightingale said her graph was designed "to affect thro' the Eyes what we fail to convey to the public through their word-proof ears."

More recent efforts by healthcare researchers like those led by Dr. Jack Wennberg at the Dartmouth Atlas Project have documented glaring—and, for the most part, inexplicable—variations in how medical resources are apportioned and delivered in the United States. The project builds on Medicare data to provide



Map 4. Percent of cancer patients admitted to intensive care during the last month of life (deaths occurring 2003-07)

Source: The Dartmouth Institute: *The Dartmouth Atlas of Healthcare*

comprehensive information and analysis about national, regional, and local markets, as well as individual hospitals and their affiliated physicians.

Consider the map reproduced from the Dartmouth Atlas Report: "Quality of End-of-Life Cancer Care for Medicare Beneficiaries." It displays the percent of cancer patients admitted to intensive care during the last month of life compared by hospital referral regions. About 24% of cancer patients nationwide were admitted to intensive care at least once during that last month. However, the percent thus admitted varied more than sevenfold across those regions (dark red versus light yellow areas on the map). This map leads the viewer to ask, "Why are these rates so dramatically different across the country?," and perhaps to add an even more significant question: "What should the rate be?"

Geospatial displays of data like this one make the variation in end-of-life care jump off the page in a way that it never would if the data were buried in a table or report narrative. Such geospatial maps and accompanying reports, coupled with the underlying research, have helped policymakers, the media, healthcare analysts, and others improve their understanding of the efficiency and effectiveness of our healthcare system. As with the map created by John Snow, the visualizations built for the Dartmouth project make the story easy to see and understand and have formed the foundation for many of the nation's ongoing efforts to improve American health and healthcare systems.

Throughout history the power of data visualizations to help us see and consider the stories buried in our health and healthcare data has resulted in profound insights and often changes and improvement in our health and healthcare. And now with the advent of technology that allows us to amass data and quickly explore it, we are on the cusp of revolutionary insights and improvements. It is an exciting time indeed.