

the truthful art



data, charts, and maps
for communication

alberto cairo

"Cairo sets the standard for how data should be understood, analyzed, and presented. *The Truthful Art* is both a manifesto and a manual for how to use data to accurately, clearly, engagingly, imaginatively, beautifully, and reliably inform the public."

Jeff Jarvis, professor, CUNY Graduate School of Journalism,
and author of *Geeks Bearing Gifts: Imagining New Futures for News*

Praise for *The Truthful Art*

“Alberto Cairo is widely acknowledged as journalism’s preeminent visualization wiz. He is also journalism’s preeminent data scholar. As newsrooms rush to embrace data journalism as a new tool—and toy—Cairo sets the standard for how data should be understood, analyzed, and presented. *The Truthful Art* is both a manifesto and a manual for how to use data to accurately, clearly, engagingly, imaginatively, beautifully, and reliably inform the public.”

—Jeff Jarvis, professor at CUNY Graduate School of Journalism and author of *Geeks Bearing Gifts: Imagining New Futures for News*

“A feast for both the eyes and mind, Alberto Cairo’s *The Truthful Art* deftly explores the science—and art—of data visualization. The book is a must-read for scientists, educators, journalists, and just about anyone who cares about how to communicate effectively in the information age.”

—Michael E. Mann, Distinguished Professor, Penn State University and author of *The Hockey Stick and the Climate Wars*

“Alberto Cairo is a great educator and an engaging storyteller. In *The Truthful Art* he takes us on a rich, informed, and well-visualized journey that depicts the process by which one scrutinizes data and represents information. The book synthesizes a lot of knowledge and carefully explains how to create effective visualizations with a focus on statistical principles. *The Truthful Art* will be incredibly useful to both practitioners and students, especially within the arts and humanities, such as those involved in data journalism and information design.”

—Isabel Meirelles, professor at OCAD University (Canada) and author of *Design for Information*

“As soon as I started immersing myself in *The Truthful Art*, I was horrified (and somewhat ashamed) to realize how much I didn’t know about data visualization. I’ve spent most of my career pursuing a more illustrative way to present data, but Alberto Cairo’s clarifying prose superbly explained the finer points of data viz. Since Alberto warns us that “[data is] always noisy, dirty, and uncertain,” everyone in this business had better read his book to find out how to properly construct visualizations that not only tell the truth, but also allow us to interact meaningfully with them.”

—Nigel Holmes, founder of Explanation Graphics

“To communicate data clearly, you have to think about it clearly. *The Truthful Art* dives deep and provides an enlightened introduction to the ‘power tools’ of data experts: science, statistics, and visualization.”

—Fernanda Viégas and Martin Wattenberg, Google

“*The Truthful Art* is essential reading for my visual communication students and for anyone (at any level) who cares about telling a story visually. Get this book, read it, act on it. If you’re looking for help to put your data visualization on the right track, this is it.”

—John Grimwade, School of Visual Communication, Ohio University

“If I were smarter, had more patience with academia, and had more focus, I might turn out to be more like Alberto, closer to the brilliance that he applies to the nature of information architecture. His title explains a lot: truth represents a most fundamental of attitudes, in questions asked, answers given and journeys taken. This [book] is a must on your thoughtful shelf of understanding.”

—Richard Saul Wurman

The Truthful Art: Data, Charts, and Maps for Communication

Alberto Cairo

New Riders

Find us on the Web at www.newriders.com

New Riders is an imprint of Peachpit, a division of Pearson Education.

To report errors, please send a note to errata@peachpit.com

Copyright © 2016 by Alberto Cairo

Acquisitions Editor: Nikki Echler McDonald

Production Editor: Tracey Croom

Development Editor: Cathy Lane

Copy Editor: Cathy Lane

Proofers: Patricia Pane, Kim Wimpsett

Compositor: Kim Scott, Bumpy Design

Indexer: James Minkin

Cover and Interior Designer: Mimi Heft

Cover Illustration: Moritz Stefaner

Notice of Rights

All rights reserved. No part of this book may be reproduced or transmitted in any form by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. For information on getting permission for reprints and excerpts, contact permissions@peachpit.com.

Notice of Liability

The information in this book is distributed on an “As Is” basis without warranty. While every precaution has been taken in the preparation of the book, neither the author nor Peachpit shall have any liability to any person or entity with respect to any loss or damage caused or alleged to be caused directly or indirectly by the instructions contained in this book or by the computer software and hardware products described in it.

Trademarks

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and Peachpit was aware of a trademark claim, the designations appear as requested by the owner of the trademark. All other product names and services identified throughout this book are used in editorial fashion only and for the benefit of such companies with no intention of infringement of the trademark. No such use, or the use of any trade name, is intended to convey endorsement or other affiliation with this book.

ISBN 13: 9780321934079

ISBN 10: 0321934075

9 8 7 6 5 4 3 2 1

Printed and bound in the United States of America

To my father

Acknowledgments

I always chuckle when someone calls me an “expert” on visualization or infographics. As a journalist, I’ve made a profession of being an amateur, in the two senses of the word: someone who doesn’t have a deep understanding of anything, but also someone who does what he does due to unabashed love for the craft.

This book is a tribute to that second kind of amateur, folks who bring good data to the world in a time when society is drowning in tsunamis of spin and misinformation. They know that it is possible to change the world for the better if we repeat the truth often and loud enough.



I’d like to first thank my University of Miami (UM) colleague Rich Beckman. It’s not an exaggeration to say that I wouldn’t be where I am today without his help, advice, and mentorship.

Seth Hamblin, a friend at *The Wall Street Journal*, passed away while I was writing this book. Seth was in love with infographics and visualization. When I told him about *The Truthful Art*, he got as excited as a kid. He was a beautiful human being, and he’ll be missed.

To Greg Shepherd, dean of UM’s School of Communication; Nick Tsinoremas, director of UM’s Center for Computational Science (CCS); Sawsan Khuri, also from CCS, a great colleague and better friend; Sam Terilli, head of our department of journalism; and Kim Grinfeder, who leads our Interactive Media program. Also at UM, I’d like to thank my colleagues in the departments of Journalism and Interactive Media, and at the Center for Communication, Culture, and Change.

To Phil Meyer, whose book *Precision Journalism* inspired me many years ago. In fact, I wrote *The Truthful Art* with the goal of being a *Precision Journalism* for the new century.

To my past colleagues at *La Voz de Galicia*, *Diario16*, DPI Comunicación, *El Mundo*, and Editora Globo. Among them, Helio Gurovitz, the former managing editor at Globo’s *Época* magazine, who combines a deep knowledge of journalism with an unusual sensibility of how to use numbers and graphics.

Also, to my clients and partners worldwide, particularly Jim Friedland, for all his support in the past two years.



Many people read this book while it was in the works. My editor, Nikki McDonald, and my copyeditor, Cathy Lane, kept an eye on me at all times, and did their best to make me meet deadlines (they failed).

Stephen Few sent me detailed notes about each chapter. Steve is both a friend and arguably my most severe critic. I've done my best to incorporate as much from his feedback as possible, but not all. I know that Steve will still disagree with some of my musings, but those disagreements can be great topics to ponder while enjoying some fine wine and cheese.

Erik Jacobsen also provided detailed feedback. His notes have been invaluable.

Three statisticians, Diego Kuonen, Heather Krause, and Jerzy Wieczorek, read the most technical chapters and made sure that I was not writing anything particularly silly. Others who commented on the book are: Andy Cotgreage, Kenneth Field, Jeff Jarvis, Scott Klein, Michael E. Mann, Isabel Meirelles, Fernanda Viégas, Martin Wattenberg, and Sisi Wei. Thank you all.

Thanks also to all the individuals and organizations who let me showcase their work in *The Truthful Art*. You're too numerous to mention, but you'll see your names in pages to come.

Some of my followers on Twitter volunteered for the last round of proofreading. They are: Mohit Chawla, Fernando Cucchietti, Stijn Debruwre, Alex Lea, Neil Richards, Frédéric Schütz, and Tom Shanley.

Special thanks to Moritz Stefaner, for giving me permission to use one of his amazing graphics on the cover of this book.



To Nancy.



Finally, and above all, thanks to my family.

About the Author

Alberto Cairo is the Knight Chair in Visual Journalism at the School of Communication of the University of Miami (UM), where he heads specializations in infographics and data visualization. He's also director of the visualization program at UM's Center for Computational Science, and Visualization Innovator in Residence at Univisión.

He is the author of the books *Infografía 2.0: Visualización interactiva de información en prensa*, published just in Spain in 2008, and *The Functional Art: An Introduction to Information Graphics and Visualization* (New Riders, 2012).

In the past two decades, Cairo has been director of infographics and visualization at news organizations in Spain and Brazil, besides consulting with companies and educational institutions in more than 20 countries. He also was a professor at the University of North Carolina-Chapel Hill between 2005 and 2009.

Cairo's personal weblog is www.thefunctionalart.com. His corporate website is www.albertocairo.com.

His Twitter handle is @albertocairo.

Additional Materials

I designed many of the charts and maps you're about to see in *The Truthful Art*, but I haven't written much about the software I used to create them. If you're interested in learning about tools, please visit my weblog, www.thefunctionalart.com, and go to the **Tutorials and Resources** section on the upper menu.

There, you will find several articles and video lessons I recorded about programs and languages like R, iNzight, and Yeeron, among others.

Contents

Preface	xiii
Introduction	1
 PART I foundations	
1 What We Talk About When We Talk About Visualization	27
To Learn More	40
2 The Five Qualities of Great Visualizations	41
The Hockey Stick Chart	43
Truthful	45
Functional	50
Beautiful	53
Insightful	59
Enlightening	60
To Learn More	65
 PART II truthful	
3 The Truth Continuum	69
Dubious Models	71
The Mind, That Clumsy Modeler	76
Why Are We So Often Mistaken?	79
Mind Bug 1: Patternicity	81
Mind Bug 2: Storytelling	82
Mind Bug 3: Confirmation	84
Truth Is Neither Absolute, Nor Relative	86
The Skills of the Educated Person	97
To Learn More	97
4 Of Conjectures and Uncertainty	99
The Scientific Stance	100
From Curiosity to Conjectures	101
Hypothesizing	105
An Aside on Variables	105

On Studies	107
Doing Experiments	110
About Uncertainty	112
To Learn More	116

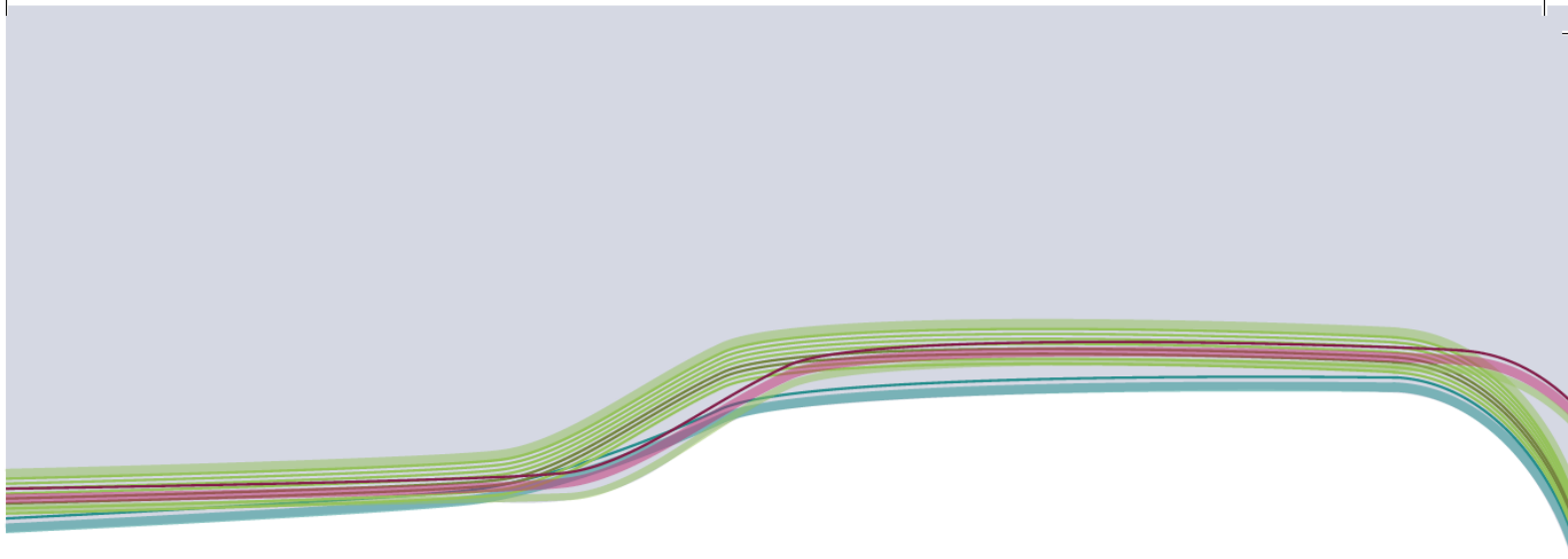
PART III functional

5 Basic Principles of Visualization	121
Visually Encoding Data	123
Choosing Graphic Forms	125
A Grain of Salt	129
Practical Tips About those Tricky Scales	135
Organizing the Display	138
Put Your Work to the Test	141
To Learn More	148
6 Exploring Data with Simple Charts	151
The Norm and the Exceptions	152
The Mode, the Median, and the Mean	153
Weighting Means	155
Range	157
Some Fun with Ranges	160
To Learn More	166
7 Visualizing Distributions	167
The Variance and the Standard Deviation	171
Standard Deviation and Standard Scores	173
Presenting Stories with Frequency Charts	178
How Standardized Scores Mislead Us	183
Peculiarities of the Normal Distribution	188
Percentiles, Quartiles, and the Box Plot	191
Manipulating Data	194
Frequency Over Time	195
To Learn More	197
8 Revealing Change	199
Trend, Seasonality, and Noise	201
Decomposing a Time Series	201
Visualizing Indexes	210

From Ratios to Logs	214
How Time Series Charts Mislead	217
Communicating Change	220
To Learn More	231
9 Seeing Relationships	233
From Association to Correlation	236
Smoothers	240
Matrices and Parallel Coordinate Plots	242
Correlation for Communication	248
Regression	257
From Correlation to Causation	259
Data Transformation	261
To Learn More	262
10 Mapping Data	263
Projection	265
Data on Maps	271
Point Data Maps	271
Line Maps	277
The Basics of Choropleth Maps	280
Classifying Data	282
Using Choropleth Maps	287
Other Kinds of Data Maps	293
To Learn More	296
11 Uncertainty and Significance	299
Back to Distributions	302
The Standard Error	305
Building Confidence Intervals	308
Disclosing Uncertainty	312
The Best Advice You'll Ever Get: Ask	316
How Has Stuff Been Measured?	317
Dispatches from the Data Jungle	318
Significance, Size, and Power	321
To Learn More	325

PART IV **practice**

12 On Creativity and Innovation	329
The Popularizers	334
News Applications	338
Those Enchanting Europeans	356
Making Science Visual	360
Culture Data	364
The Artistic Fringes	367
 Epilogue	 375
Index	376



Preface

It All Begins with a Spark

Why is it that when one man builds a wall, the next man immediately needs to know what's on the other side?

—Tyrion Lannister in George R.R. Martin's *A Game of Thrones*

There's probably something you don't know about college professors: we tend to have peculiar hobbies.

In October 2014, I spent my entire fall recess catching up with R, a programming language for statistical analysis; ggplot2, an R library that creates nice-looking charts; and Tableau, a data visualization program.¹ Learning any software tool without using it is impossible, so I needed some data to play with, and not just any data, but data I could care about.

A few months back, my family and I had moved to a new home, so I had briefly visited the Miami-Dade County Public Schools website (DadeSchools.net) to check the quality of the elementary school, middle school, and high school in our area. Each had a grade of A. I had felt reassured at the time, but also a bit

¹ I hope that this doesn't impress you. I am by no means an advanced user of any of these tools. All graphics in these pages were designed with very little knowledge of how to use them properly. For more information, visit <http://www.r-project.org/>, <http://ggplot2.org/>, and <http://www.tableau.com>.

Region	SchoolName	Reading2012	Reading2013	ReadingDifference	Math2012	Math2013	MathDifference	SchoolGrade	BoardDistrict
5	0041 AIR BASE ELEMENTAR	82	80	-2	71	75	4	A	9
7	0070 CORAL REEF MONT AC	71	73	2	64	56	-8	A	9
4	0071 EUGENIA B THOMAS K	69	69	0	66	64	-2	A	5
7	0072 SUMMERSVILLE ADVANT	57	50	-7	50	54	4	B	9
6	0073 MANDARIN LAKES K-8	34	32	-2	38	39	1	C	9
6	0081 LENORA BRAYNON SIMI	28	29	1	26	47	21	F	2
1	0091 BOB GRAHAM EDUCATI	68	70	2	68	66	-2	A	4
1	0092 NORMAN S EDELCLIP	73	72	-1	78	77	-1	A	3
7	0100 RUTER ACADEMY	68	68	0	73	76	3	A	4
4	0101 ARCOLA LAKE ELEMEN	59	52	-7	41	39	-2	C	2
7	0102 MIAMI COMMUNITY CH	38	41	3	43	47	4	D	9
4	0111 MAYA ANGELOU ELEME	45	35	-10	59	50	-9	B	5
4	0121 AUBURNDALE ELEMENT	53	51	-2	56	55	-1	A	6
4	0122 DR ROLANDO ESPINOS	65	64	-1	66	63	-3	A	5
5	0125 NORMA BUTLER BOSSA	70	67	-3	74	70	-4	A	7
5	0161 AVOCADO ELEMENTARY	45	33	-12	45	45	0		9
4	0201 BANYAN ELEMENTARY	73	74	1	72	70	-2	A	8
5	0211 DR MANUEL C BARREI	71	71	0	74	68	-6	A	7
1	0231 AVENTURA WATERWAYS	68	68	0	67	67	0	A	3
1	0241 R K BROAD/BAY HARB	76	75	-1	81	77	-4	A	3
5	0251 ETHEL ROGER BECKHA	83	80	-5	89	90	1	A	8
6	0261 BEL-AIRE ELEMENTAR	32	32	0	36	48	12	D	9
5	0271 BENT TREE ELEMENTA	70	61	-9	69	60	-9	A	8
5	0311 GOULDS ELEMENTARY	36	40	4	51	52	1	B	9
7	0312 WATER GARDENS ACAD	75	76	1	84	85	1	A	4
1	0321 BISCAYNE ELEMENTAR	45	42	-3	52	50	-2	B	3
7	0332 SOMERSET ACAD -SIL	62	66	4	54	64	10	A	9
7	0339 SOMERSET ACAD -SO	67	57	-10	60	54	-6	B	9
1	0341 ARCH CREEK ELEMENT	47	48	1	47	48	1	B	1
7	0342 PINECREST ACADEMY	72	75	3	78	76	-2	A	7
6	0361 BISCAYNE GARDENS E	37	35	-2	39	37	-2	D	1
7	0400 RENAISSANCE ELEM C	80	82	2	76	82	6	A	5

Figure P.1 The top portion of a spreadsheet with data from public schools in Miami-Dade County.

uneasy, as I hadn't done any comparison with schools in other neighborhoods. Perhaps my learning R and Tableau could be the perfect opportunity to do so.

DadeSchools.net has a neat data section, so I visited it and downloaded a spreadsheet of performance scores from all schools in the county. You can see a small portion of it—the spreadsheet is 461 rows tall—in **Figure P.1**. The figures in the Reading2012 and Reading2013 columns are the percentage of students from each school who attained a reading level considered as satisfactory in those two consecutive years. Math2012 and Math2013 correspond to the percentage of students who were deemed reasonably numerate for their age.

While learning how to write childish simple scripts in R, I created rankings and bar charts to compare all schools. I didn't get any striking insight out of this exercise, although I ascertained that the three public schools in our neighborhood are decent indeed. My job was done, but I didn't stop there. I played a bit more.

I made R generate a scatter plot (**Figure P.2**). Each dot is one school. The position on the X-axis is the percentage of students who read at their proper level in 2013. The Y-axis is the same percentage for math proficiency. Both variables

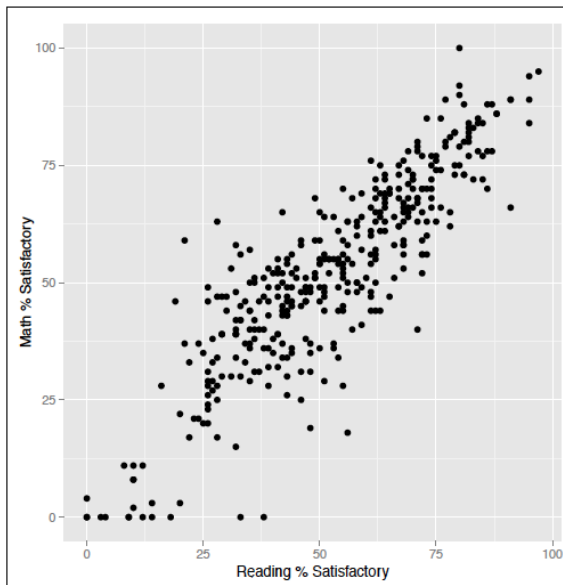


Figure P.2 Each dot on the chart is a school. Reading and math skills are strongly related.

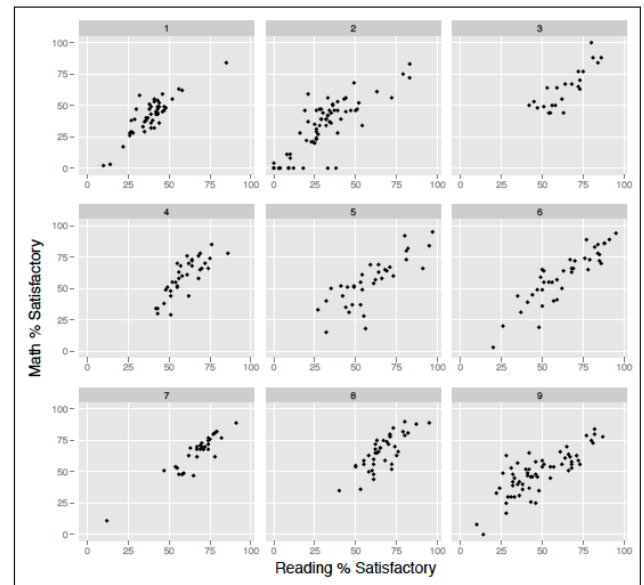


Figure P.3 The same data, divided by school board.

are clearly linked: the larger one gets, the larger the other one tends to become.² This makes sense. There is nothing very surprising other than a few outliers, and the fact that there are some schools in which no student is considered proficient in reading and/or math. This could be due to mistakes in the data set, of course.

After that, I learned how to write a short script to design not just one but several scatter plots, one for each of the nine school board districts in Miami-Dade County. It was then that I became really intrigued. See the results in **Figure P.3**.

There are quite a few interesting facts in that array. For instance, most schools in Districts 3, 7, and 8 are fine. Students in Districts 1 and 2, on the other hand, perform rather poorly.

At the time I was not familiar with the geography of the Miami-Dade school system, so I went online to find a map of it. I also visited the Census Bureau website to get a map of income data. I redesigned and overlaid them. (See **Figure P.4**. Warning: I didn't make any adjustment to these maps, so the overlap isn't perfect.) I got what I foresaw: the worst-performing districts, 1 and 2, encompass low-income neighborhoods, like Liberty City, Little Haiti, and Overtown.

² In statistics, we may call this a "strong positive correlation." But I'm getting a bit ahead of myself.

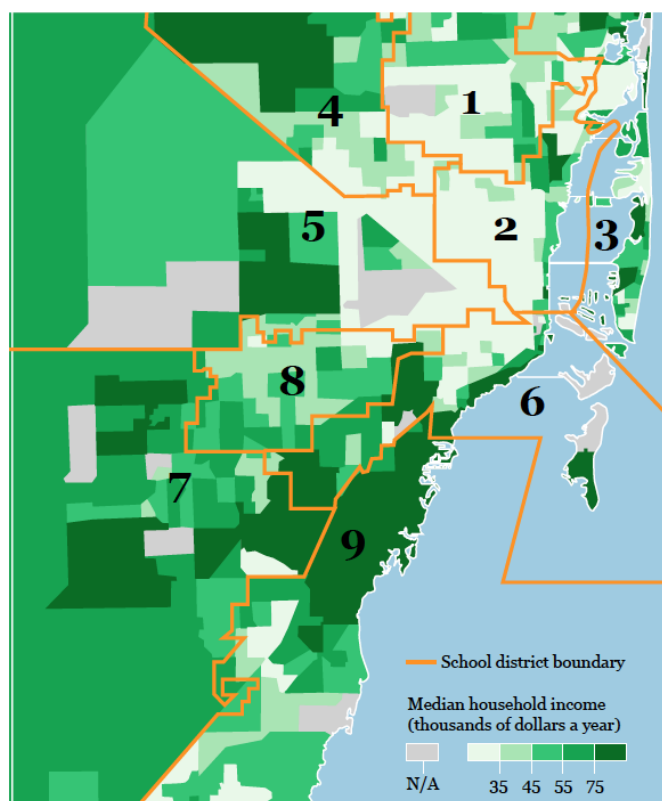


Figure P.4 Median household income in the nine school board districts of Miami-Dade County.

Immediately, questions started piling up in my head. Is the relationship between bad schools and low household income direct? Does a bad education lead to reduced wages? Or do kids coming from low-income families go to school being already disadvantaged, and that worsens the scores of the schools they attend? Am I getting causality right? What are other possible variables that affect both school performance and income?

What about the outliers in those charts, those schools in Districts 1 and 7, for instance, that are so far from their flocks? Or that school in District 3 that got a perfect score in math? And what about District 6? Schools in that plot are much more spread out than in the others. Is that related to the sharp divide between a richer strip on the east (Coconut Grove) and poorer blocks on the west within that school district?

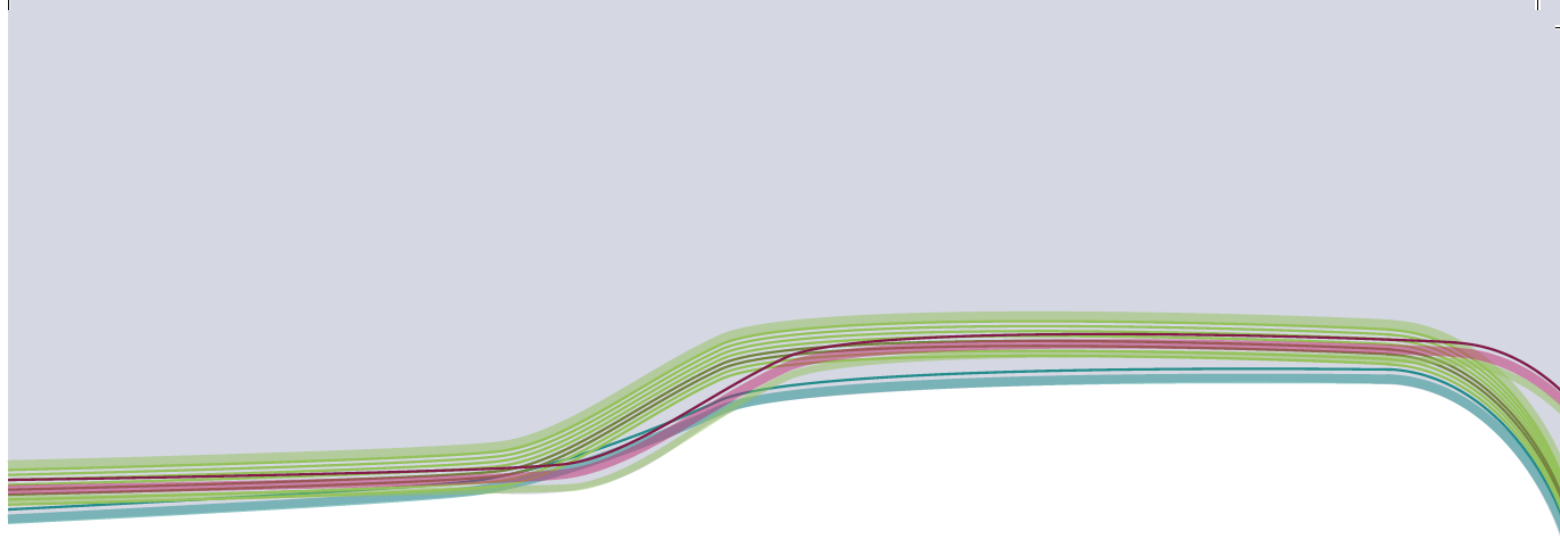
And more: have all these percentages and grades changed substantially in the past few years? If so, is it due to real variation in the quality of our public education or because of changes in the methods researchers use to measure attainment? So many questions.

And so the seeds for many potential stories got planted. I didn't have an idea of what they might be at that point or if any of them would be worth telling. I just got a glimpse, an enticing clue. As most visualization designers and data journalists I know will tell you, sometimes it is not you who finds good ideas when you're seeking them. Instead, good ideas find you in the most unexpected circumstances.

Good ideas are fleeting things, so I feverishly scribbled notes in a computer application called Stickies, short messages for my future self, musings of a mind in a state of joyous flow. I added, "Find some education experts.³ Ask them. Contact the folks running dadeschools.net. You'll likely need more data from the U.S. Census Bureau's website." And so on and so forth.

As the saying goes, every great story begins with a spark. Fun ensues.

³ Here's Robert B. Reich—who isn't an expert on education but was Secretary of Labor under President Bill Clinton—in his book *Saving Capitalism* (2015): "A large portion of the money to support public schools comes from local property taxes. The federal government provides only about 10 percent of all funding, and the states provide 45 percent, on average. The rest is raised locally (...) Real estate markets in lower-income communities remain weak, so local tax revenues are down. As we segregate by income into different communities, schools in lower-income areas have fewer resources than ever. The result is widening disparities in funding per pupil, to the direct disadvantage of poor kids." Another possible clue to follow.



Introduction

The Island of Knowledge and the Shoreline of Wonder

Never before have we had so many tools to learn and to communicate. Yet the art of talking, listening, and ascertaining the truth seems more elusive than ever in this Internet and cable age, lost in a bitter stream of blather and misinformation.

—Maureen Dowd, “Toilet-Paper Barricades,”
The New York Times (August 11, 2009)

It’s a fact of life that kids put their parents to the test with their boundless and unpredictable curiosity. I have three small children, ages 4, 8, and 10. Here are some of the questions that I got from them while writing these lines, followed by my thoughts and the actual answers I gave them.

Ten-year-old boy: “Daddy, how can I build a hobbit house in the Minecraft video game?”

My thought: “That’s a damn good question. Would it be possible to build J.R.R. Tolkien’s Middle Earth in Minecraft?” (Sorry, I’m a nerd.)

My answer: “No idea, kiddo. I don’t even know how to play Minecraft. Why don’t you watch *The Lord of the Rings* movies again? I will be happy to watch them with you.”

Four-year-old girl: “Daddy, do you have a baby in your tummy?”

Thought: “Man, kids can be brutally honest.”

Answer: “No, sweetie. Only mommies can carry babies in their tummies. Daddy just has big muscles.”

Eight-year-old girl: “Daddy, why do planets never stop spinning?”

Thought: “Huh...”

Answer: “Why don’t you Google it, honey?” (Just kidding, although I believe that many parents in my position would go for that one.)

My actual answer was: “Can you give me a few hours? Then I’ll be able to explain it to you.”

My daughter’s question forced me to take my fingers off the keyboard and think for a minute. My memory of Newtonian physics was a bit rusty, but I was sure that the fact that planets keep rotating is related to the laws of motion. I dusted a couple of popular science books from my shelves and also looked for some articles on the Internet. Then I grabbed a pencil, a pen, and some crayons: I think more clearly when I draw. I ended up with a series of sketches that I’m going to call an “infographic,” a graphical display intended to convey information.

Here’s the story I told my daughter.

Our First Infographic

Let’s begin here, on planet Earth. When you throw a ball (**Figure I.1**), it tends to move forward and spin around its own center. The faster we throw the ball and the heavier it is, the more momentum, or impetus, it carries. There are two kinds of momentum in this case: linear (forward motion) and angular (spinning.)

It is obvious that the ball will not move forever. Eventually, it will stop. Why? First of all, the ground, as well as the air, provides friction (**Figure I.2**). Air is a

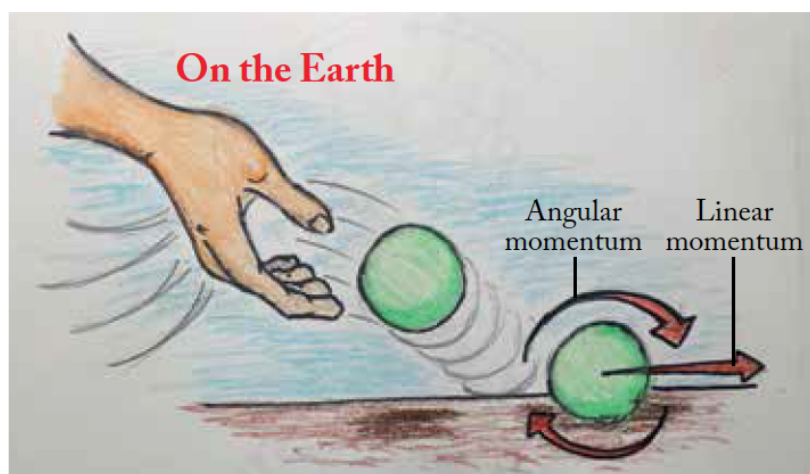


Figure I.1 The first sketch that I made for my daughter. When you throw a ball to the floor, it tends to move forward (linear momentum) and to rotate around its own axis (angular momentum).

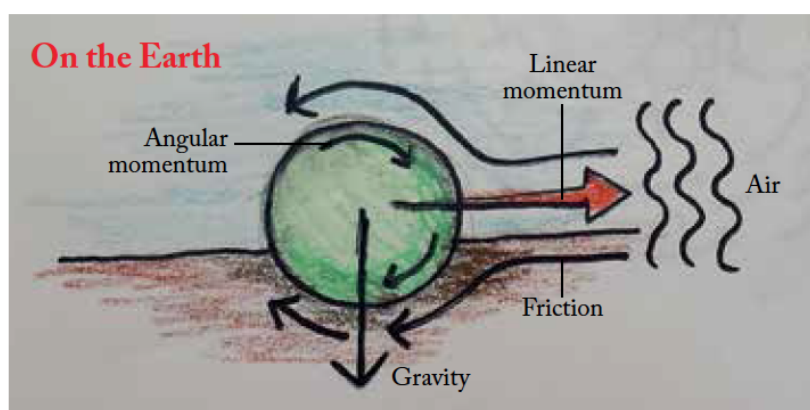


Figure I.2 Gravity and friction with the air and with the ground will make the ball stop moving.

fluid, like water. Imagine that you jump into a swimming pool. When you enter the water, your speed decreases, right? That's friction at work. Friction is a word we use to describe an interesting phenomenon: the air and the ground absorb the momentum that the ball carries.

Now, imagine that you're an astronaut floating in deep space (**Figure I.3**). (My daughter pointed out that this illustration is inaccurate; the person in it should be wearing a spacesuit! Extra credit for her.) There's no air in space. Therefore, there's almost no friction. If we throw our imaginary green ball, it is likely that

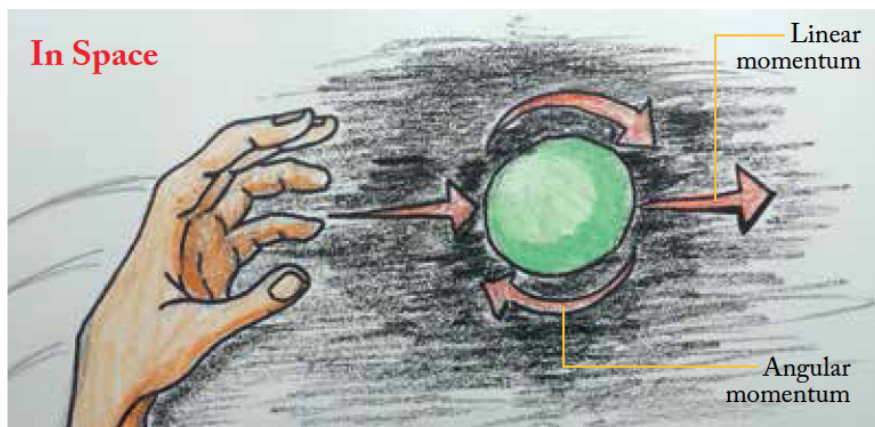


Figure I.3 No visual explanation is perfect, and this one isn't an exception. Notice that the drawing isn't accurate, and not just because the person in it isn't wearing a space suit. If the hand adopts that position when throwing, the ball won't spin that way, or at all.

it will take millions and millions of years to stop moving forward. It won't stop spinning until then either. We call this **conservation of momentum**: if there's nothing to interfere with the ball, it won't stop moving.

Next, let's go back in time zillions of years, to the era when none of the planets in the solar system existed. Just the sun was there, surrounded by large clouds of dust particles. These clouds spun at a very, veeeery slow speed.¹ The particles were held together and bound to the sun by gravity (**Figure I.4**).

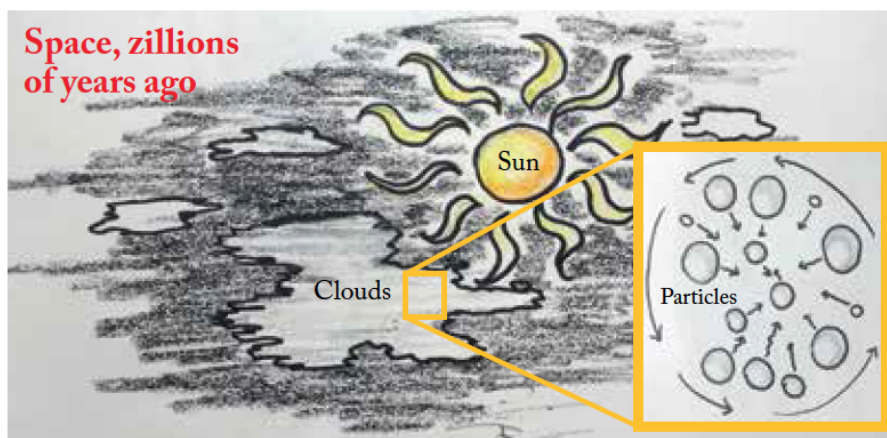


Figure I.4 Our solar system, zillions of years ago.

¹ Where did this initial momentum come from? There's some homework for you!

Something interesting happened then: little by little, gravity made these little particles move closer and closer to each other. And the closer they were, the faster they spun. The reason why that happened is a bit complicated to explain, but it isn't crucial for now, so let's leave it for another day.

The last step is the easiest to understand. The particles in the clouds moved so close to each other that they ended up merging. Earth and the other planets were born from collapsing dust clouds (**Figure I.5**), and they keep spinning, at least for now, because there's very little in space to stop them from doing so.

I must confess that my daughter didn't get all this at first. It was just a bit too much information to digest. So I went over the entire sequence of drawings with her again. While doing it, I realized that this exercise embodied a few points that I make in all my courses, namely:

- **When you design a graphic to explain something, getting the information right comes first.** No good infographic or data visualization—we'll learn the difference between the two soon—can be based on deficient data and analysis. The quality of your graphics depends fundamentally on the quality of your reporting or research, not just on how good a graphic designer you are.
- **Being concise and clear doesn't imply oversimplifying.** Any act of communication involves a controlled reduction of complexity, up to the point when reducing matters further would hurt the integrity of the information. I adapted the message to my audience by getting rid of jargon and equations, but I respected the essence of the facts and my daughter's intelligence.
- **Good design isn't about embellishment but about structuring information to enable understanding.** That said, aesthetic appeal is a worthy goal, as it can help make our messages more attractive and, as a consequence, more effective.
- **Graphics that encode information function as cognitive aids.** If I had described the process using *just* words, it's likely that you'd feel compelled to mentally visualize little balls and arrows. My drawings are intended to do that work for your brain.

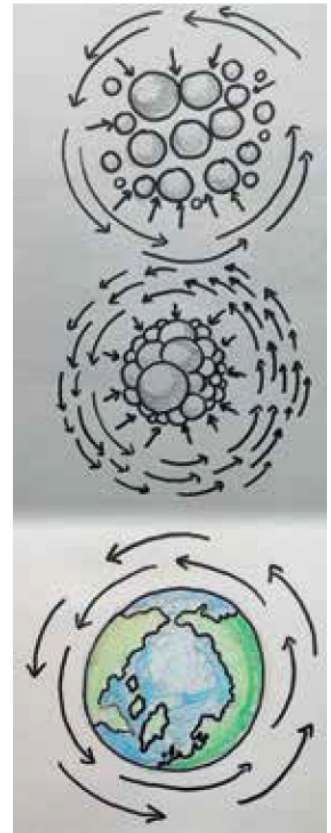


Figure I.5 The particles collapse and form the planets in the solar system.

- **If words are sometimes useless by themselves, so are charts, maps, diagrams, and illustrations.** It is in the *combination* of words (spoken or written) and visuals that the magic of understanding often happens.

Quite good for such a simple exercise, right? And there's more.

The Island of Knowledge

I got another insight from this tale about spinning planets. After I finished the explanation for the second time, my daughter remained silent for a few seconds. Then she asked, "OK, I get it. But Earth spins very fast. Why aren't we thrown into space, then?"

Understanding never quenches the thirst for *more* understanding, does it? Quite the contrary is true: the more we learn, the more aware we become of the gaps in our knowledge. As physicist **Marcelo Gleiser** wrote, "The knowledge that we have defines the knowledge that we can have... As knowledge shifts, we ask new kinds of questions that we couldn't have anticipated."²

Gleiser's book is titled after the most evocative metaphor I know about the human quest for understanding: the island of knowledge. It seems that the first person to refer to it was New York Methodist pastor **Ralph W. Sockman**, who is believed to have said: "**The larger the island of knowledge, the longer the shoreline of wonder.**"³

Chet Raymo (1998) expanded the metaphor beautifully:

All scientific knowledge that we have of this world, or will ever have, is as an island in the sea of mystery. We live in our partial knowledge as the Dutch live on polders claimed from the sea. We dike and fill. We dredge up soil from the bed of mystery and build ourselves room to grow.

Of course, being a visualization designer, I couldn't resist making a little illustration of what happened inside my daughter's brain (**Figure I.6**): the island of knowledge expanded for her, but so did the shoreline of wonder, draining new land out of the sea of mystery.

² In his book *The Island of Knowledge*, (2014).

³ I haven't been able to track the source of this quote, so take this with a grain of salt.

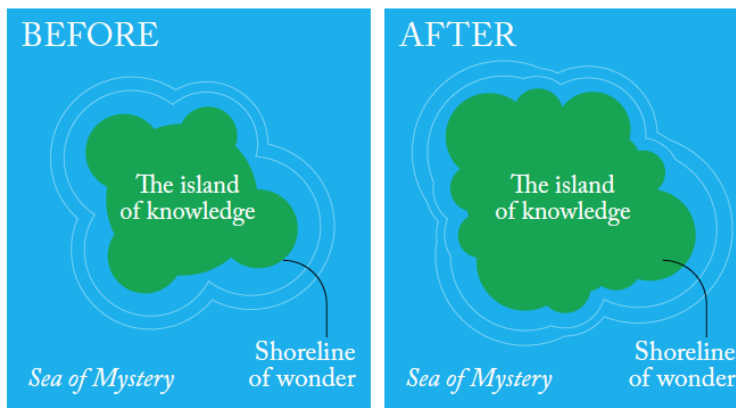


Figure I.6 The island of knowledge and the shoreline of wonder.

Finding the right answers to good questions makes us capable of posing even better and more profound ones. We can all enjoy this process. If you allow me a brief digression, this is the reason why I believe that we should teach our kids to love not knowledge per se, but *learning*. They should be encouraged to cherish the very quests they undertake, not just their products.

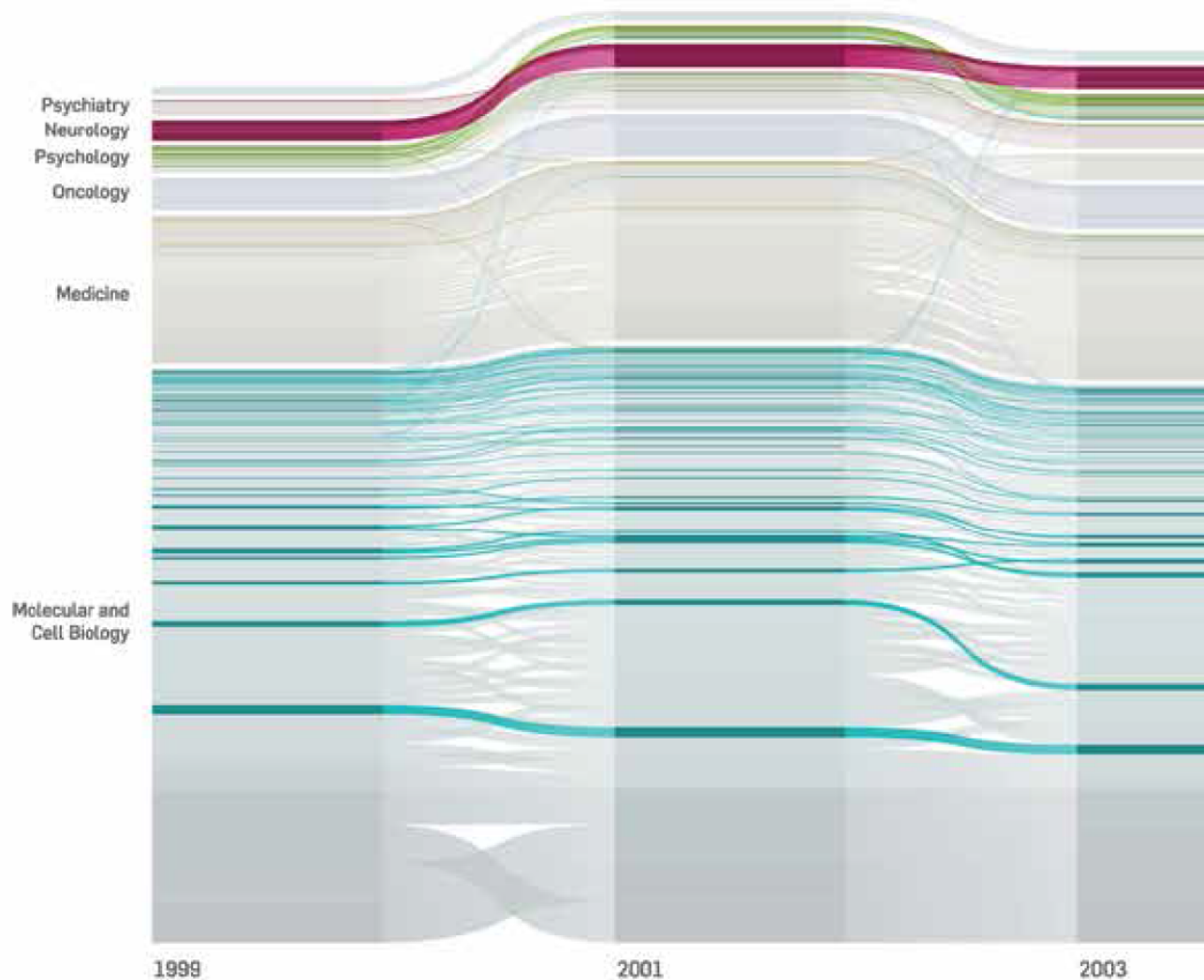
Good data visualizations and explanation infographics **communicate information** and, as a result, they can increase our understanding. That's their first role, and it's what I was focusing on when I made the drawings for my daughter.

But **graphics may also prompt exploration**. They reveal as much as much as they lead to new questions. A graphic may tell you a compelling story, but it can also invite you to expand the shoreline of wonder in ways that its creators didn't predict. Here's famous visualization designer **Moritz Stefaner** in a manifesto titled "Worlds, not stories."⁴

Data visualization can help us both to understand complex issues a bit better, but also to provide images to debate about, to refer back to, and sometimes just to meditate over... I want [people] to use the visualizations I provide as starting points for their own explorations... Consequently, any serious visualization of a sufficiently complex topic should always aim at exposing the complexity, the inner contradictions, the manifold nature of the underlying phenomenon. I like to provide users with a structured way to explore a complex phenomenon on their own terms, in a sensually rich mosaic of media and facts rather than a pre-digested narrative with a surprise at the end. To me, interesting topics rarely boil down to a single story.

⁴ Moritz Stefaner: <http://well-formed-data.net/archives/1027/worlds-not-stories>

The emergence of neuroscience

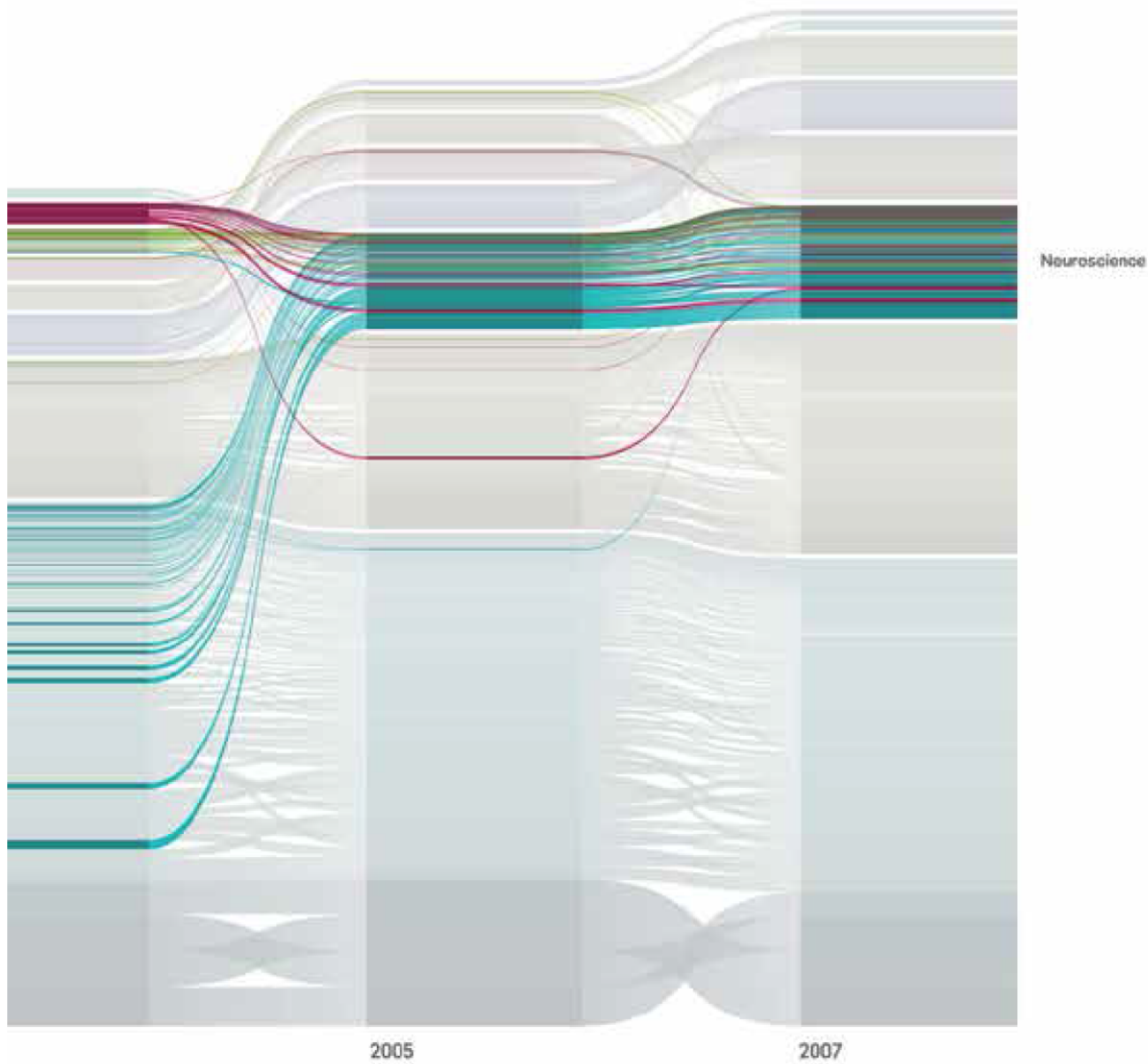


This visualization documents the formation of neuroscience as a field of its own right over the last decade. Originally scattered across related disciplines (such as medicine, molecular and cell biology or neurology), the neuroscientific journals start to define a niche of their own, reflected in the dense cluster emerging in 2005.

First, almost 8000 scientific journals are clustered into groups, based on citation patterns, and using the map equation. In short, for a network, the map equation specifies the theoretical limit of how compactly a network can be described by a trajectory of a random walker on the network. Thereof, the map equation over all possible network partitions reveals regular information flow across directed and weighted networks or, in our case, the flow of how citations flow through science.

Second, using the Eigenfactor™ Score, the journals are assigned a measure of importance – much as Google's PageRank algorithm ranks the importance of web pages. The Eigenfactor™ Score measures the percentage of time that a random walker would spend with the respective journal, if they were to move through the network by randomly following citations in the journals.

Figure 1.7 Visualization by Moritz Stefaner (<http://moritz.stefaner.eu>), in collaboration with Martin Rosvall, Jevin West, and Carl Bergstrom at the Bergstrom Lab, University of Washington.



ups, based on their
a network partitioned
t of how concisely we
rk. Therefore, minimiz-
eals regularities of
r, in our case, the struc-

igned a measure of im-
o importance of web
f time that researchers
ve through the network

This process is repeated in two-year chunks from 1999–2007, in order to capture changes in clustering and shifts in importance over the years. For this diagram, we picked only the clusters relevant to the formation of neuroscience.

In the visualization, each cluster occupies a vertical column block in the respective year's column, further subdivided into a block for each journal. Each journal is connected with a horizontal band over the years. The height of each journal reflects the Eigenfactor Score. All journals in the cluster that corresponds to the field of neuroscience in year 2007 are highlighted to tell the story of the formation of this field of science. The coloring is based on the cluster assignments in the first year, 1999.

Visualization: Moritz Stefaner (<http://moritz.stefaner.eu>)
Data analysis: Eigenfactor team (<http://eigenfactor.org>)

<http://well-formed.eigenfactor.org>

Stefaner doesn't reject traditional linear narrative techniques outright, but he prefers to build interactive displays that enable discovery. His work embodies this approach. In 2009, he partnered up with the Bergstrom lab in the biology department at the University of Washington to create **Eigenfactor**,⁵ a project that visualizes citation patterns between scientific journals.

Some of those visualizations are illuminating. Take a look at **Figure I.7** (on the previous pages). What do you notice? To me, besides the striking beauty of this Sankey diagram,⁶ its central message is clear: modern neuroscience is the result of the confluence of several disciplines.

That's the main story I extract. Yours can be different than mine. What you'll get from this graphic depends on the knowledge that you had before facing it. As a journalist, I can be standing on a different place of the shoreline of wonder than a scientist. Stefaner's visualization may expand the shoreline in slightly different directions for each of us (**Figure I.8**) and, therefore, it will lead us to stare at two increasingly disparate patches of the horizon above the sea of mystery.

A few weeks ago, I stumbled upon a graphic designed by data scientist **Gilad Lotan** (**Figure I.9**) that further illustrates how the communication and exploration

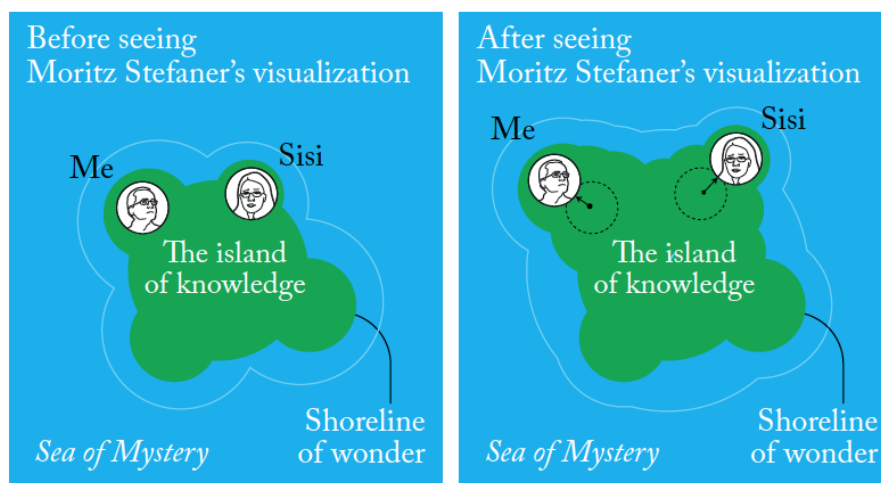


Figure I.8 The island of knowledge expands in different directions, depending on your previous knowledge. The role of scientist in this illustration is played by my friend Sisi Wei, who works for the nonprofit investigative journalism organization ProPublica.

5 Visualizing information flow in science:
<http://truth-and-beauty.net/projects/well-formed-eigenfactor/>

6 There's some jargon for you! We'll learn about Sankey diagrams soon.

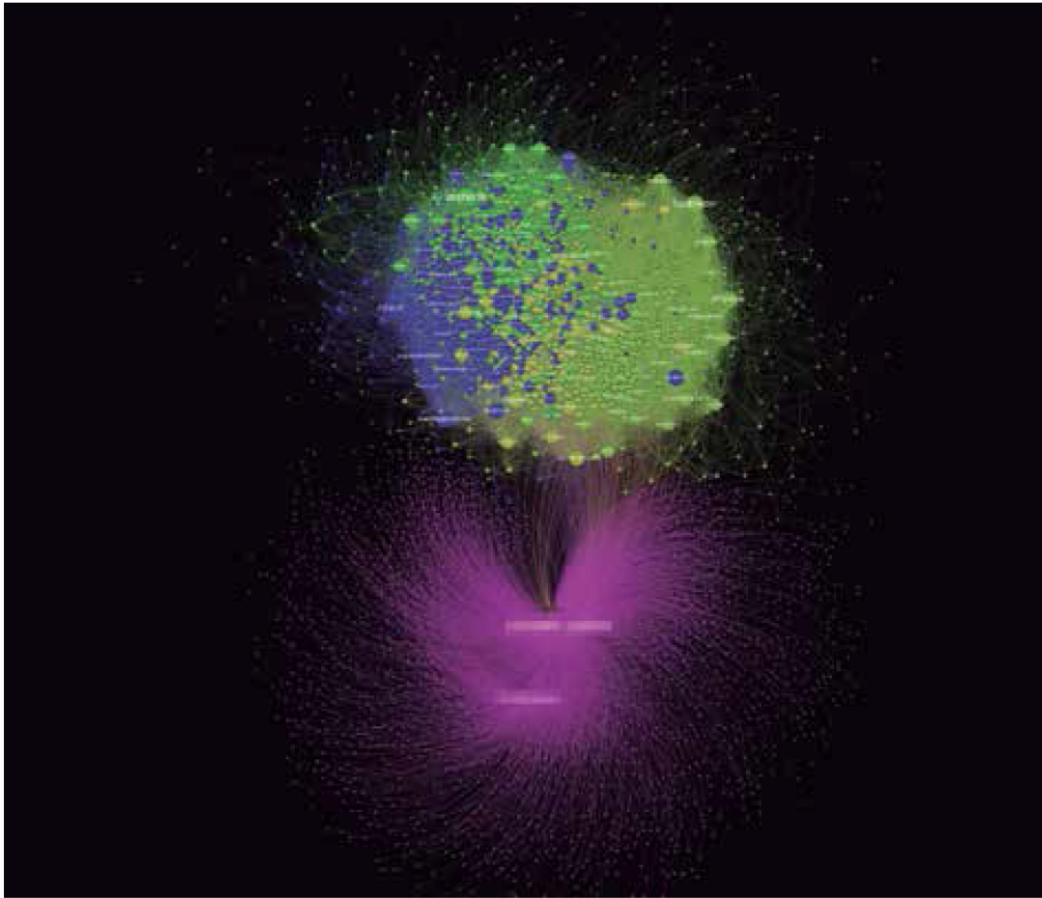


Figure I.9 Visualization by Gilad Lotan (<http://giladlotan.com/>). Lotan writes, “This is a network graph representation of my Twitter followers after I acquired the bots. The top cluster represents my ‘real’ followers, who are intertwined: many follow each other, clearly a community of users. The bottom purple region represents the fake accounts, who are completely separate—structurally they’re clearly not a real community, with very little connectivity between the accounts.”

dimensions of visualization design complement each other. In an article published in the *Los Angeles Times* and on the website Medium.com,⁷ Lotan described a particularly quirky experiment.

In the past few years, services that promise to increase your standing on social media are thriving. For a few bucks, you can be followed by thousands of fake Twitter accounts. Lotan wanted to answer two questions: first, does buying fake followers lead to more real followers? And second, how are fake followers

⁷ Gilad Lotan: ‘(Fake) friends with (Real) benefits’:
<https://medium.com/i-data/fake-friends-with-real-benefits-eec8c4693bd3>

connected to accounts created by actual people? The short answer to the first question is yes. To understand why, you should read Lotan's article.

The answer to the second question is in the visualization that Lotan designed with the data he collected. The multi-colored blob on top represents his 2,600 followers before the experiment. Being part of a virtual community, the nodes are strongly connected and close to each other. The purple cloud at the bottom corresponds to the 4,000 faux followers Lotan paid for. Notice how scattered and sparsely linked they are. No community exists here. This is the main message revealed by transforming thousands of data points into a visual shape.

Many more insights could hide in this intricate display, though. To discover them, we'd need to scrutinize it attentively. Again, communication and exploration go hand in hand. We can tell stories with graphics, but we can also let people build their own stories with them.

To rephrase all this a bit, let's say that a good visualization is:

1. reliable information,
2. visually encoded so relevant patterns become noticeable,
3. organized in a way that enables at least some exploration, when it's appropriate,
4. and presented in an attractive manner, but always remembering that honesty, clarity, and depth come first.

This, in a nutshell, is what half of this book is about.

The other half is about what *precedes* design. There are obstacles that hinder the expansion of the island of knowledge. It's not possible to become a professional visualization designer without learning how to overcome them.

Some of these obstacles arise from the territoriality of academic disciplines and the lack of communication between them. I still remember the first time that I read an introduction to cartography textbook, for instance, and my surprise at realizing that many of the design principles that I had learned in journalism school were almost identical to the ones that makers of data maps follow. Academically speaking, I'm one of the least territorial people you'll ever meet, so in this book I will shamelessly borrow from graphic design, journalism, the philosophy of science, statistics, cartography, and many other areas.

Other obstacles to the expansion of the island of knowledge are much more insidious.

Candid Versus Strategic Communication

This is a book not just about how to design information graphics but about how to design *candid* information graphics. At its core lies a simple idea: **The purpose of infographics and data visualizations is to enlighten people—not to entertain them, not to sell them products, services, or ideas, but to inform them.** It's as simple—and as complicated—as that.

What I call candid communication is practiced by professionals whose main goal is (or *should* be) to increase society's collective knowledge. They usually come from disciplines like the sciences, statistics, journalism, cartography, information design, and so on. Their importance in a world increasingly brimming with nonsense and propaganda is paramount.

This is no trivial challenge. See **Figure I.10**. It compares the number of public relations specialists with the number of journalists in the United States. It

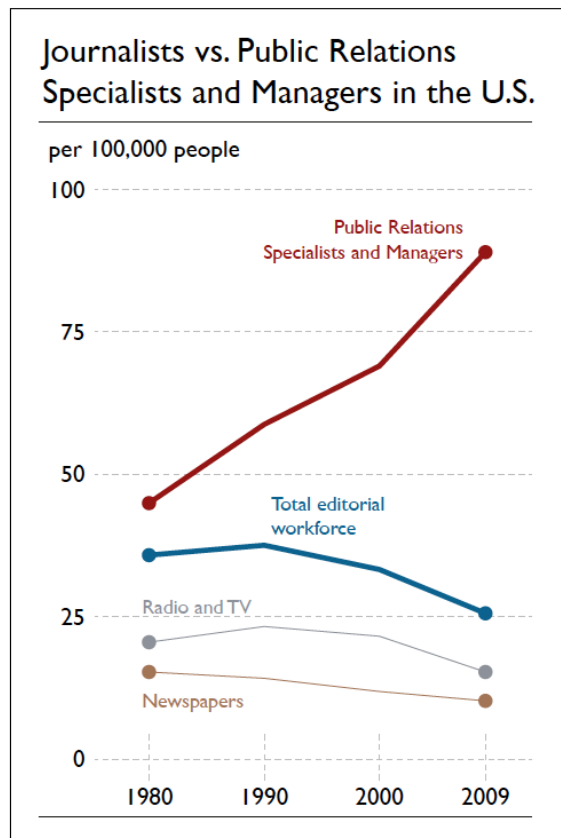


Figure I.10 The number of professionals working in public relations has expanded greatly in the past three decades, while the number of journalists has dropped. (Graph based on McChesney and Nichols, 2011.)

reminds me of a problem I've written about elsewhere:⁸ in the past few years, strategic communication—the world of advertisement, PR, marketing, etc.—has hijacked the very word “infographics.” This word used to define data-rich graphical displays intended to spread newsworthy information. It has a long and noble history in the journalism industry.⁹

Today, the word “infographics” generally means puerile posters used as clickbait. Search for the word on the Internet and you'll see what I mean. No candid communication here, in general. You'll mostly find bland, simplistic, and tendentious visuals based on shaky data, designed mainly to draw Web traffic, not to inform.

Needless to say, journalists can and do lie to the public, and most strategic communicators may be honest, but it's also fair to acknowledge that “convey your best understanding of what the truth is” is the core mandate for journalists, while “never deceive the public [but] present the facts in a way that sheds as much positive light [on your cause or company] as possible”¹⁰ is the central position for specialists in strategic communication. The second portion of that sentence is crucial, as it renders the whole statement contradictory.

One of the most intriguing books I've read recently is *All Marketers Are Liars* (2005). The title is a joke—in part. Its author, **Seth Godin**, writes that marketing is about creating compelling stories. He explains, “A great story is true. Not true because it's factual but true because it's consistent and authentic.”

Godin excuses himself saying that marketers “are just storytellers. It's the consumers who are liars. As consumers, we lie to ourselves every day... Successful marketers are just the providers of stories that consumers choose to believe.”

As a good marketer, Godin enjoys toying with words. I'll have some things to say in this book about storytelling in visualization, but I can give you a heads-up: the stories I'm interested in are not the ones that are shiny simulacra of truth, no matter how “authentic” and “consistent” they are. In fact, reality is rarely

8 Reclaiming the word “infographics,”

<http://www.thefunctionalart.com/2012/12/claiming-word-infographics-back.html>

9 The most venerable news graphics conference is called the “Malofiej Infographics Summit” for a reason: <http://www.malofiejgraphics.com/>.

10 Principles of Public Relations:

<http://smallbusiness.chron.com/10-principles-public-relations-10661.html>. A classic definition of PR: “Public relations is a strategic communication process that builds mutually beneficial relationships between organizations and their publics” (1982, Public Relations Association of America's National Assembly). This is as vague as it gets. That “mutually” doesn't mean that the relationship is *equally* beneficial for the organizations and for the public.

consistent, but messy and nuanced. Therefore, messiness and nuance should be part of any candid act of graphics communication.

Second, there's a substantial difference between *lying* and *deceiving* (ourselves or others). Lying is always a conscious action. You know what the reality is, but you choose to frame it in a way that advances your agenda. Deceit, though, isn't necessarily conscious. You can simply be oblivious of reality and, as a consequence, you may mislead yourself or others.

Third, it's true that human beings can't be completely factual or objective. Our brain is a flawed meat machine chiseled by evolution, not a computer. We all have cognitive, cultural, and ideological biases, but that doesn't mean that we can't strive to be factual. *Truth* is unattainable, but **trying to be truthful is a realistic and worthy goal**, and there are certain techniques that can help us pursue it.

There's a deep difference between those who surrender to their own biases, or willingly embrace them, and those who work hard to identify and curb them, even if they'll never completely succeed. This book is a tribute to this second group.

Continuing with Godin: In another part of his bestseller he writes, "Maybe who is lying to whom isn't all that important, in the end, as long as the connection has been made and the story has been successfully told." In other words, what really matters is if consumers buy your product, service, or idea, not if any of them has any substance.

Godin points out that marketing techniques can be used to advance good causes. No doubt about that. But I'd like to add something right away: **in strategic communication, you may begin with a message and then look for information to support it. In candid communication, you begin with the information, and then you thoroughly analyze it to discover the messages worth spreading are.** Or you begin with a message, but then you collect your information in a way that it could lead you to refute that very message.

If our quest for truthfulness leads us to conclude that the stories we wish to tell or the graphics we are so eager to design are inaccurate or plainly untrue—well, we'll need to be ready to drop them outright. We must never distort our data so they fit into our preconceived narratives, no matter how much we love them.

You may think that this is just a platitude. Twisting reality a bit to get your story across effectively isn't that inappropriate, is it? After all, we all do it every day, and nobody gets hurt. Nobody? It's all a matter of degree.

Let's think of an extreme case. The **Discovery Institute**, a non-profit organization based in Seattle, is dedicated to misinforming people about science. This is not really how the folks who work there describe themselves, of course, but it's exactly what they do. Numerous books by scientists, like Donald R. Prothero's *Evolution: What the Fossils Say and Why it Matters* (2007), include long sections describing Discovery's misdeeds in detail.

Discovery's main activity is to popularize creationism, although the organization prefers to call it "intelligent design." This is the fallacy that Darwinian evolution cannot account for the appearance of new species in our planet.

Discovery has a budget of millions of dollars. Most of it seems to be devoted to promotion, as the organization doesn't conduct any proper research. Its websites, graphics, books, and lectures are sleek and lavishly produced. Discovery does fantastic marketing and PR—according to Godin's definitions—but its marketing and PR hurt society by distorting data and spreading baloney.

Excellent strategic communication, which is what Discovery excels at, isn't always backed up by excellent information. If you think that this doesn't have harmful consequences, think again. According to the Pew Research Center, 33 percent of U.S. adults are incorrectly convinced that "humans existed in present form since the beginning of time."¹¹

It's rare to find a case that is as clear-cut as this one. Truth and untruth aren't absolutes but stand at the end points of a fuzzy continuum. Still, a warning is worthwhile: **there are people out there who aren't in the business of expanding the island of knowledge by navigating beyond the shoreline of wonder. They are in the business of transforming that shoreline into a dark, impassable marsh.**

In the kind of visualizations I'm interested in, the quality of the information precedes the quality and visual appeal of the graphics themselves. It's impossible to have one without the other. Many pundits, marketers, and activists of all ideological stripes are willfully ignorant of this precept. This endangers us all by muddling our public discourse.

This is why the chart in Figure I.10 concerns me so much. I feel uncomfortable with the world it announces.

¹¹ www.pewforum.org/2013/12/30/publics-views-on-human-evolution/

In the Good Ol' Days...

Before the advent of the World Wide Web, in the early 1990s, we used to get our information mainly from newspapers, radio, and TV. If you are my age, you probably remember the pre-WWW days, waiting for the morning paper or the evening news to learn what had happened in your community, your country, or around the world.

News organizations, and the journalists who worked for them, acted as gatekeepers: they decided what information was worth publishing. They chose their sources—sometimes unwisely—and they filtered and shaped our view of reality.

It was far from an ideal world. I'm no romantic. Journalists are as prone to error and bias as anyone else. We can be tricked by spin doctors, and we often miss the most relevant stories of our time.

Not to mention that news organizations have carelessly helped promote bad ideas, such as climate change denialism and unproven alternative medical therapies, just to name a couple.¹²

It's hardly surprising that trust in news media and professional journalists is low today. But robust journalism serves us well. Here is **Charles Lewis**, a famous investigative reporter:

In a society increasingly beset by public relations, advertising, and other artificial sweeteners manufactured by message consultants and communications flacks, how does an ordinary citizen decipher truth...?¹³

Helping people tell bullshit from facts should certainly be a duty for all journalists and information designers. Whenever we get information from anyone, we have to make sure that it's reliable, to the extent of our knowledge and analytical skills. We need to ask our sources: **how do you know?** And moreover: **how do you know that you know?**, as that will force them to disclose the methods and data used to reach their conclusions.

¹² Supporters of these causes have traditionally been very fond of professional marketing, by the way. Do read Ben Goldacre's *Bad Science: Quacks, Hacks, and Big Pharma Flacks* (2010) when you have three or four hours to spare. You're welcome.

¹³ See "To Learn More" at the end of this chapter.

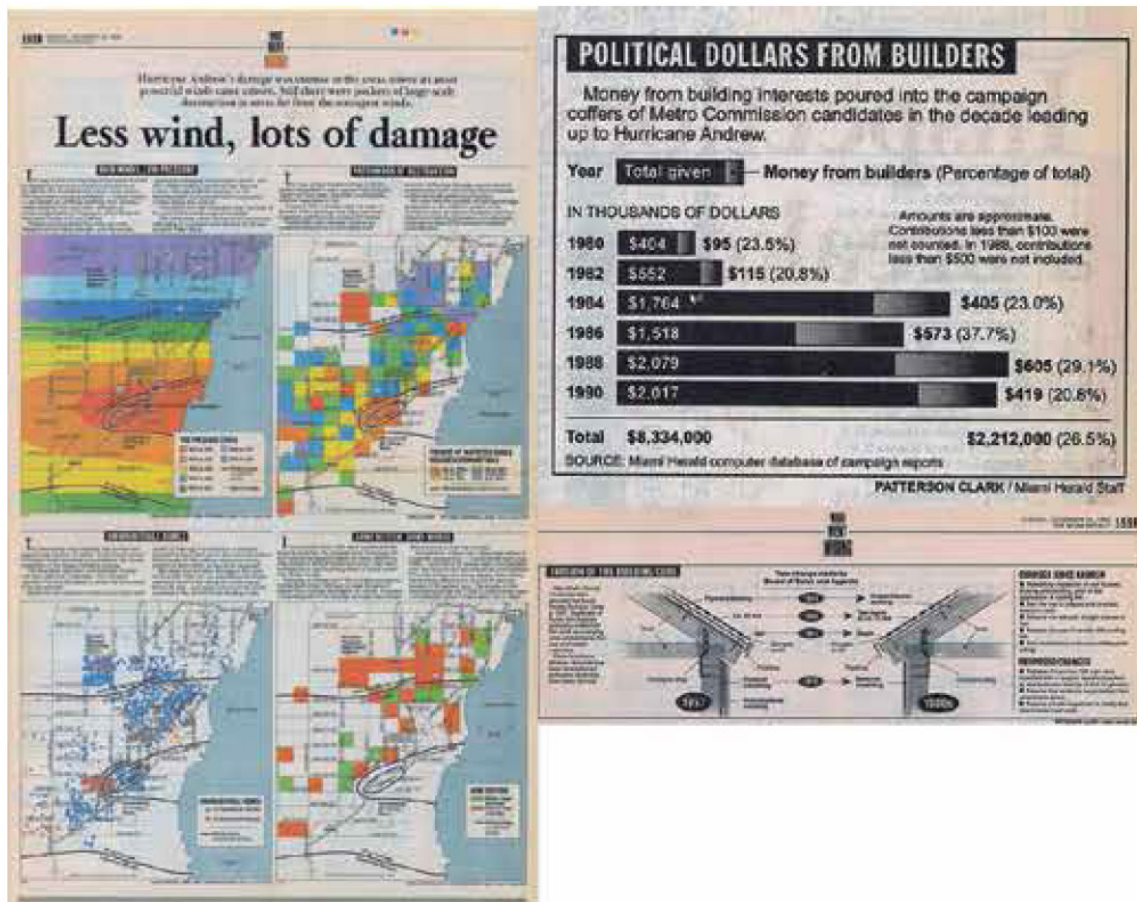
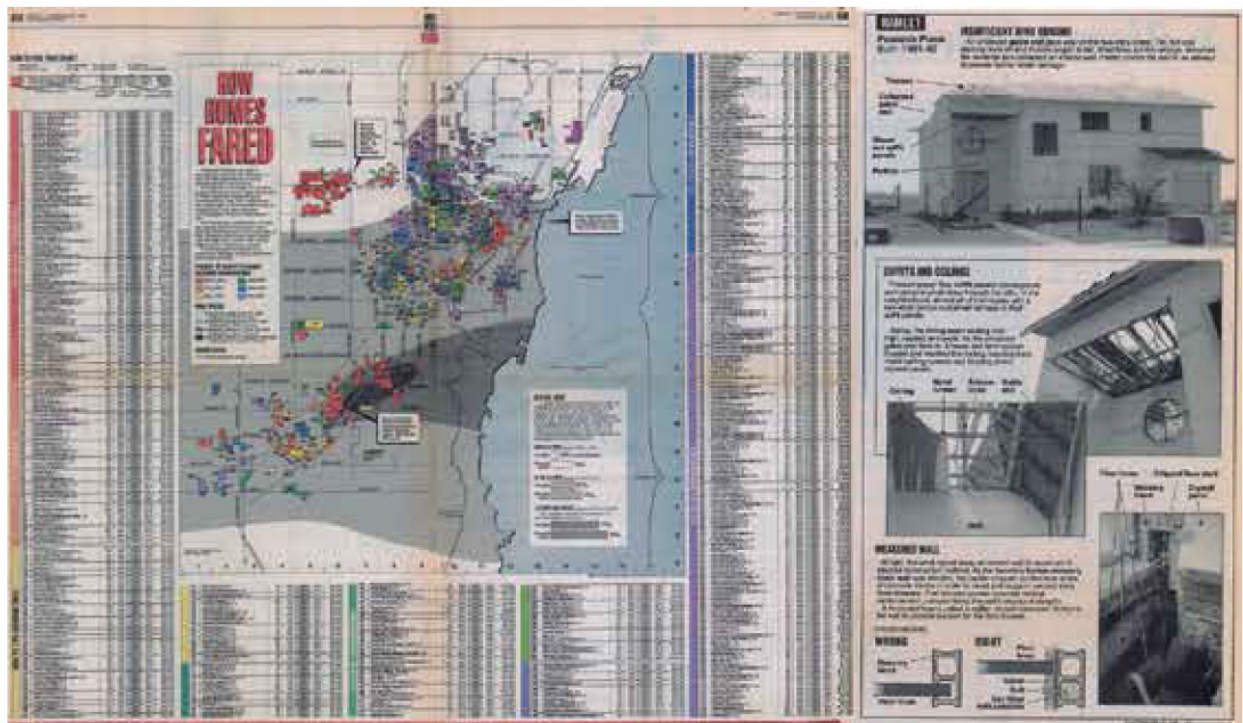
Journalists also do original reporting about relevant matters. Here by my side I have the August 24, 2014, edition of the *Miami Herald*. One long story in this issue explains how Haitian Americans in Florida are being preyed upon by Ponzi scheme crooks who advertise in Creole-language radio. Another one is a well-balanced study of a violent Liberty City neighborhood in Miami.

The *Herald* has a long tradition of heavy-hitting investigative reporting. In August 1992, category 5 Hurricane Andrew leveled much of south Miami-Dade county. In December that year, the *Herald* published a special 16-page supplement titled “What Went Wrong” that revealed that houses built between the 1980s and the 1990s were more likely to have suffered serious damage than older ones, due to lax inspection of building quality and zoning in the decade that preceded the hurricane. The supplement was supported by plenty of graphics (**Figure I.11**), and it won a Pulitzer Prize for public service in 1993.

When I see stories like these, I wonder: aren’t we in danger of losing a central pillar of a democratic society, now that so many traditional news organizations are at risk of crumbling? If the *Miami Herald* and other publications like it disappear, who is going to expose wrongdoing in a systematic manner so the public can be aware of it?

You could argue that plenty of startups populate the current media landscape: Vox.com, Mic.com, FiveThirtyEight, The Huffington Post, Gawker, Quartz, BuzzFeed, and so on. Aren’t they protecting their readers against the roaring tides of noise and spin? Even if they try to, not all of them have been very effective so far. According to what **Ryan Holiday** says in his scathing jeremiad *Trust Me, I’m Lying: Confessions of a Media Manipulator* (2012), many online news organizations don’t fact-check their sources, don’t do proper editing, and don’t produce much original reporting.

Figure I.11 (opposite) A sample of the graphics published by the *Miami Herald* on December 20, 1992. Here’s how the *Herald* described the project: “IN TODAY’S *Herald* is a 16-page report titled ‘What Went Wrong.’ Investigative reporters working with engineers and other experts examine why so many homes that shouldn’t have suffered heavy destruction during Hurricane Andrew did (...) The first comprehensive analysis of Andrew’s havoc concludes that shoddy construction helped turn a devastating storm into the most costly disaster in U.S. history. During a four-month investigation, the *Herald* analyzed damage reports on 60,000 houses and matched these by computer with millions of property and building records.” (Reporting, charts, maps, and illustrations by Stephen K. Doig, Jacquee Petchel, Dan Clifford, Lisa Getter, Patterson Clark, Jeff Leen, Luis Soto, and Don Finefrock.)



There are exceptions, of course. Vice News, for example, is deservedly famous for its excellent and edgy video documentaries. And, to be fair, the companies mentioned before are investing more resources every year in improving their journalism.

In the meantime, **what's left for regular citizens to do?** Now that anyone can reach hundreds, thousands, or even millions of people through personal websites, blogs, and social media, who is going to make sure that information that goes viral on Twitter, Facebook, or on any other platform is accurate? And who's going to create original visualizations that help people stay informed on important matters, and conduct better lives?

Perhaps the answer is me. And you. Any and all of us.

Your Inner Skeptic, Your Inner Journalist

Here's a little secret: most of the maps, charts, and diagrams that I'll praise in this book are "journalistic" in the sense that the designers who created them first did their best to make sure that their information was relevant, factual, and accurate, and only then did they present it in a way that was accessible and engaging.

However, not all of those designers call themselves journalists.

What Is a Journalist, Anyway?

When I was in college, I remember that one of my professors said that a journalist is someone in charge of producing the news. And what is "the news"? Here we can borrow from **Jack Fuller**, ex-president and publisher of the *Chicago Tribune*: "News is a report of what a news organization has recently learned about matters of some significance or interest to the specific community that news organization serves."

Got it. But here's a story: back in 2010, when I was the director of infographics and multimedia at the Brazilian weekly news magazine *Época*, I heard of a young computer scientist called **Maurício Maia**. Maurício lives in São Paulo, a city that suffers from heavy rains during the summer. Many streets in the city flood periodically, so he decided to create an interactive map and database of

past floods, to identify the best and the worst areas. He downloaded publicly available datasets from government websites, and he transformed them into a tool that he made available for free.¹⁴ What would you call what Maurício did? I call it journalism.

I'm not the only one. **Jeff Jarvis**, a professor at City University of New York, once wrote that "there are no journalists, there is only the service of journalism."¹⁵ Maybe this is a bit too extreme, but it rings true. If you are a designer, you may have recalled **Laszlo Moholy-Nagy's** famous saying, "**Designing is not a profession but an attitude.**" We are all designers, in the sense that we're creatures who have a taste for organizing matter and ideas into objects and patterns. Perhaps we all are—or can become—journalists, too, every once in a while at least.

The Elements of Journalism, a classic introduction to the practice, says that "the purpose of journalism is to provide people with the information they need to be free and self-governing." It then lists what is needed to fulfill this task:

1. Journalism's first obligation is to the truth.
2. Its first loyalty is to citizens.
3. Its essence is a discipline of verification.
4. Its practitioners must maintain an independence from those they cover.
5. It must serve as an independent monitor of power.
6. It must provide a forum for public criticism and compromise.
7. It must strive to make the significant interesting and relevant.
8. It must keep the news comprehensive and proportional.
9. Its practitioners must be allowed to exercise their personal conscience.

I'd like to argue that these shouldn't be the exclusive values of a specific professional group. Society as a whole will be better off if they become civic values that the entire population embraces. The world will turn into a much nicer place if more of us learn about what data and evidence are and

¹⁴ Maurício was puzzled when I phoned him to say that I wanted to hire him to work on my team. He said, "I am not a journalist!" I didn't succeed; he was making too much money as a freelancer. His website, in which he discusses his projects, is <http://mmaia.tumblr.com/>

¹⁵ Jeff Jarvis: "There are no journalists"
<http://buzzmachine.com/2013/06/30/there-are-no-journalists-there-is-only-journalism/>

become more critical as a result. And it'll be even nicer if we also learn how to convey that evidence in a clear, compelling, and useful manner through data visualizations, infographics, or interactive and searchable tables.

Maurício Maia committed an act of journalism. He spent time devising a visual tool that citizens of São Paulo could take advantage of. I believe that acts of candid, evidence-based communication and of useful design like Maia's aren't frequent enough, while marketing, PR, and advocacy are pervasive. Nothing against any of the latter—they have a role in a market economy—but the former are the ones that I hold dear and will feature in this book.

Who is this book for, then? This book is written first for designers and journalists who wish to communicate effectively with data visualization and infographics. If you're in this group, this book may give you a glimpse into the worlds of science, statistics, and information design. It won't teach you everything you need to know, but it may open many doors.

I also write for anyone who isn't a professional designer or journalist, but wishes to understand and use visualization. Therefore, this book is also for scientists, data analysts, business intelligence types, etc.

If you are in this second group, please take a look at another quote by **Charles Lewis**:

Imagine a world in which individual researchers [and many other kinds of experts] and journalists are sometimes looking for truth in all the same places, using the same exciting new data technologies and analytics, exchanging ideas and information, and sometimes working and writing together, whether side by side or across borders and genres. These collaborative fact-finders, fact-checkers, truth-seekers, and truth-tellers will all come from different perspectives.... but all will share the deep curiosity, patience, determination, and mettle that have always characterized the investigative reporter.¹⁶

Let me suggest something similar: we live in a world where special interests spend millions of dollars pushing their agendas, promoting their ideologies, and selling their unsubstantiated claims with almost no interference. They can reach us more effectively than in any previous era thanks to the Internet and social media. They like to use data, infographics, and visualization because they believe that people trust them more than they trust mere words.

¹⁶ LEWIS (2014), page 224

Arguably, it's time to push back. Perhaps after reading this book you will consider joining the virtual alliance of people whose purpose is to expand the island of knowledge by exploring the mysteries that lie beyond the shoreline of wonder. This process is much simpler than you may think. You decide that you have something relevant to communicate; you gather data and scrutinize them thoroughly to make sure that you get everything right; and then you design your graphics.

I will be delighted if, ultimately, I am able to convince you that this is an endeavor worth pursuing.

To Learn More

Authors who claim ownership of everything they write are delusional. Therefore, at the end of each chapter you will find a section like this that will lead you to the readings that inspired most of my thoughts. There's also a complete bibliography in the last pages.

Here are some books to expand on the content of this introduction, besides the ones mentioned in the text itself:

- Gleiser, Marcelo. *The Island of Knowledge: The Limits of Science and the Search for Meaning*. New York: Basic Books, 2014.
- Goldacre, Ben. *Bad Science: Quacks, Hacks, and Big Pharma Flacks*. New York: Faber and Faber, 2010.
- Lewis, Charles. *935 Lies: The Future of Truth and the Decline of America's Moral Integrity*. New York: PublicAffairs, 2014.
- McChesney, Robert Waterman, and John Nichols. *The Death and Life of American Journalism: The Media Revolution that Will Begin the World Again*. Philadelphia, PA: Nation, 2010.
- Raymo, Chet. *Skeptics and True Believers: The Exhilarating Connection between Science and Religion*. New York: Walker, 1998.
- Starkman, Dean. *The Watchdog That Didn't Bark: The Financial Crisis and the Disappearance of Investigative Journalism*. New York: Columbia University Press, 2014.