

---

# Fundamentals of Data Visualization

*A Primer on Making Informative  
and Compelling Figures*

*Claus O. Wilke*

Beijing • Boston • Farnham • Sebastopol • Tokyo

**O'REILLY®**

---

# Table of Contents

<b>Preface.....</b>	<b>xi</b>
<b>1. Introduction.....</b>	<b>1</b>
Ugly, Bad, and Wrong Figures	2
<hr/>	
<b>Part I. From Data to Visualization</b>	
<b>2. Visualizing Data: Mapping Data onto Aesthetics.....</b>	<b>7</b>
Aesthetics and Types of Data	7
Scales Map Data Values onto Aesthetics	10
<b>3. Coordinate Systems and Axes.....</b>	<b>13</b>
Cartesian Coordinates	13
Nonlinear Axes	16
Coordinate Systems with Curved Axes	22
<b>4. Color Scales.....</b>	<b>27</b>
Color as a Tool to Distinguish	27
Color to Represent Data Values	29
Color as a Tool to Highlight	33
<b>5. Directory of Visualizations.....</b>	<b>37</b>
Amounts	37
Distributions	38
Proportions	39
x-y relationships	41
Geospatial Data	42

Uncertainty	43
<b>6. Visualizing Amounts</b>	<b>45</b>
Bar Plots	45
Grouped and Stacked Bars	50
Dot Plots and Heatmaps	53
<b>7. Visualizing Distributions: Histograms and Density Plots</b>	<b>59</b>
Visualizing a Single Distribution	59
Visualizing Multiple Distributions at the Same Time	64
<b>8. Visualizing Distributions:</b>	
<b>Empirical Cumulative Distribution Functions and Q-Q Plots</b>	<b>71</b>
Empirical Cumulative Distribution Functions	71
Highly Skewed Distributions	74
Quantile-Quantile Plots	78
<b>9. Visualizing Many Distributions at Once</b>	<b>81</b>
Visualizing Distributions Along the Vertical Axis	81
Visualizing Distributions Along the Horizontal Axis	88
<b>10. Visualizing Proportions</b>	<b>93</b>
A Case for Pie Charts	93
A Case for Side-by-Side Bars	97
A Case for Stacked Bars and Stacked Densities	99
Visualizing Proportions Separately as Parts of the Total	101
<b>11. Visualizing Nested Proportions</b>	<b>105</b>
Nested Proportions Gone Wrong	105
Mosaic Plots and Treemaps	107
Nested Pies	111
Parallel Sets	113
<b>12. Visualizing Associations Among Two or More Quantitative Variables</b>	<b>117</b>
Scatterplots	117
Correlograms	121
Dimension Reduction	124
Paired Data	127
<b>13. Visualizing Time Series and Other Functions of an Independent Variable</b>	<b>131</b>
Individual Time Series	131
Multiple Time Series and Dose–Response Curves	135

Time Series of Two or More Response Variables	138
<b>14. Visualizing Trends</b> .....	<b>145</b>
Smoothing	145
Showing Trends with a Defined Functional Form	151
Detrending and Time-Series Decomposition	155
<b>15. Visualizing Geospatial Data</b> .....	<b>161</b>
Projections	161
Layers	169
Choropleth Mapping	172
Cartograms	176
<b>16. Visualizing Uncertainty</b> .....	<b>181</b>
Framing Probabilities as Frequencies	181
Visualizing the Uncertainty of Point Estimates	186
Visualizing the Uncertainty of Curve Fits	197
Hypothetical Outcome Plots	201

---

## Part II. Principles of Figure Design

<b>17. The Principle of Proportional Ink</b> .....	<b>207</b>
Visualizations Along Linear Axes	208
Visualizations Along Logarithmic Axes	212
Direct Area Visualizations	215
<b>18. Handling Overlapping Points</b> .....	<b>219</b>
Partial Transparency and Jittering	219
2D Histograms	222
Contour Lines	225
<b>19. Common Pitfalls of Color Use</b> .....	<b>233</b>
Encoding Too Much or Irrelevant Information	233
Using Nonmonotonic Color Scales to Encode Data Values	237
Not Designing for Color-Vision Deficiency	238
<b>20. Redundant Coding</b> .....	<b>243</b>
Designing Legends with Redundant Coding	243
Designing Figures Without Legends	250

<b>21. Multipanel Figures.....</b>	<b>255</b>
Small Multiples	255
Compound Figures	260
<b>22. Titles, Captions, and Tables.....</b>	<b>267</b>
Figure Titles and Captions	267
Axis and Legend Titles	270
Tables	273
<b>23. Balance the Data and the Context.....</b>	<b>277</b>
Providing the Appropriate Amount of Context	277
Background Grids	282
Paired Data	287
Summary	290
<b>24. Use Larger Axis Labels.....</b>	<b>291</b>
<b>25. Avoid Line Drawings.....</b>	<b>297</b>
<b>26. Don't Go 3D.....</b>	<b>305</b>
Avoid Gratuitous 3D	305
Avoid 3D Position Scales	307
Appropriate Use of 3D Visualizations	313

---

## Part III. Miscellaneous Topics

<b>27. Understanding the Most Commonly Used Image File Formats.....</b>	<b>319</b>
Bitmap and Vector Graphics	319
Lossless and Lossy Compression of Bitmap Graphics	321
Converting Between Image Formats	324
<b>28. Choosing the Right Visualization Software.....</b>	<b>325</b>
Reproducibility and Repeatability	326
Data Exploration Versus Data Presentation	327
Separation of Content and Design	330
<b>29. Telling a Story and Making a Point.....</b>	<b>333</b>
What Is a Story?	334
Make a Figure for the Generals	337
Build Up Toward Complex Figures	341

Make Your Figures Memorable	343
Be Consistent but Don't Be Repetitive	345
<b>Annotated Bibliography.....</b>	<b>351</b>
<b>Technical Notes.....</b>	<b>355</b>
<b>References.....</b>	<b>357</b>
<b>Index.....</b>	<b>361</b>