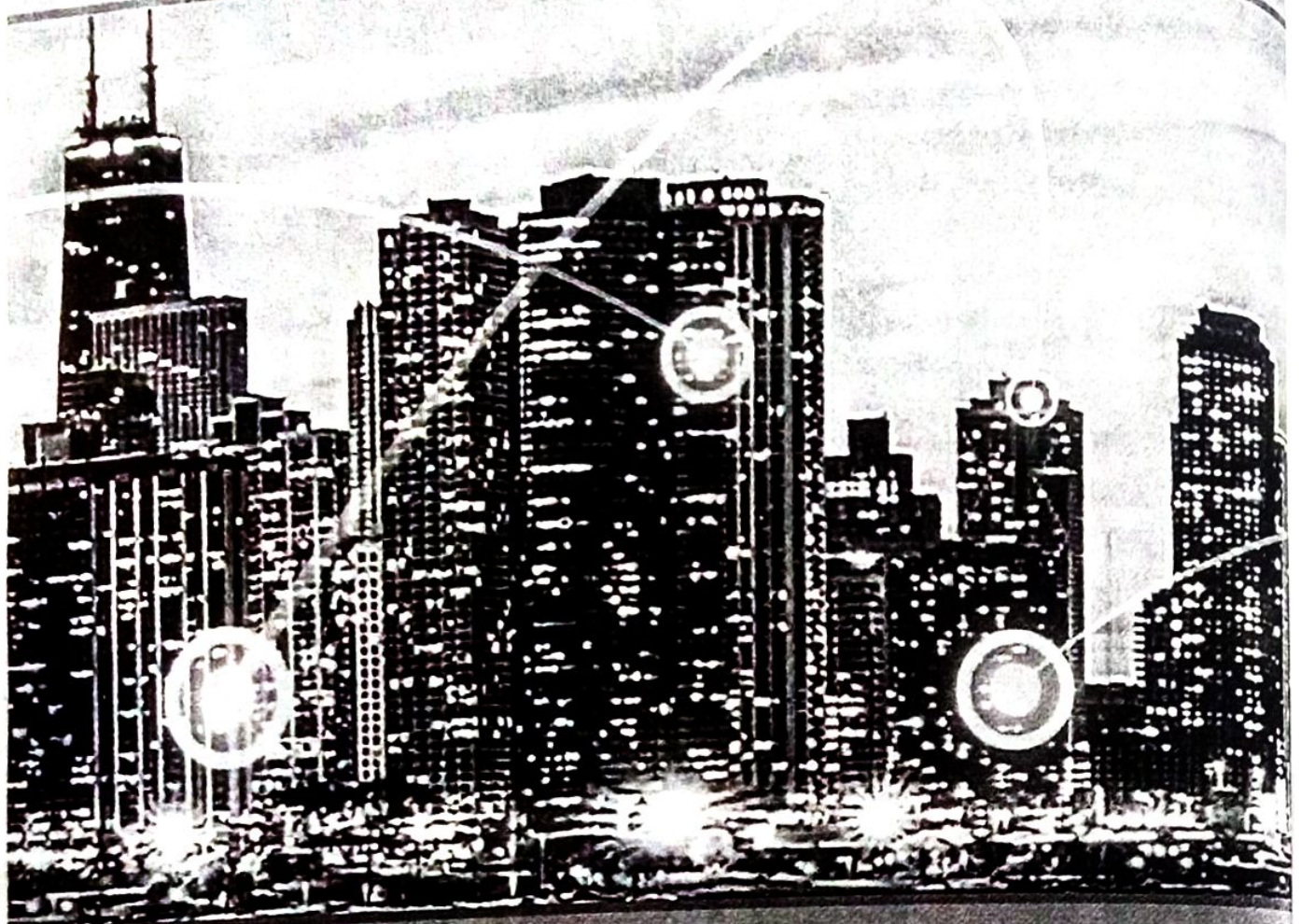




Data Visualization

Exploring and Explaining with Data

Jeffrey D. Camm
Wake Forest University

Michael J. Fry
University of Cincinnati

James J. Cochran
University of Alabama

Jeffrey W. Ohlmann
University of Iowa



Australia • Brazil • Canada • Mexico • Singapore • United Kingdom • United States

Copyright 2012 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has determined that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Contents

ABOUT THE AUTHORS xii

PREFACE xiii

CHAPTER 1 Introduction 2

1.1 Analytics 3

1.2 Why Visualize Data? 4

Data Visualization for Exploration 4

Data Visualization for Explanation 7

1.3 Types of Data 8

Quantitative and Categorical Data 8

Cross-Sectional and Time Series Data 9

Big Data 10

1.4 Data Visualization in Practice 11

Accounting 11

Finance 12

Human Resource Management 13

Marketing 14

Operations 14

Engineering 16

Sciences 16

Sports 17

Summary 18

Glossary 19

Problems 20

CHAPTER 2 Selecting a Chart Type 26

2.1 Defining the Goal of Your Data Visualization 28

Selecting an Appropriate Chart 28

2.2 Creating and Editing Charts in Excel 29

Creating a Chart in Excel 30

Editing a Chart in Excel 30

2.3 Scatter Charts and Bubble Charts 32

Scatter Charts 32

Bubble Charts 33

2.4 Line Charts, Column Charts, and Bar Charts 35

Line Charts 35

Column Charts 39

Bar Charts 41

2.5 Maps 42

Geographic Maps 42

Heat Maps 44

Treemaps 45

2.6	When to Use Tables	47
	Tables versus Charts	47
2.7	Other Specialized Charts	49
	Waterfall Charts	49
	Stock Charts	51
	Funnel Charts	52
2.8	A Summary Guide to Chart Selection	54
	Guidelines for Selecting a Chart	54
	Some Charts to Avoid	55
	Excel's Recommended Charts Tool	57
	Summary	59
	Glossary	60
	Problems	61
CHAPTER 3	Data Visualization and Design	76
3.1	Preattentive Attributes	78
	Color	81
	Form	81
	Length and Width	84
	Spatial Positioning	87
	Involvement	87
3.2	Gestalt Principles	88
	Similarity	88
	Proximity	88
	Enclosure	89
	Connection	89
3.3	Data-Ink Ratio	91
3.4	Other Data Visualization Design Issues	98
	Minimizing Eye Travel	98
	Choosing a Font for Text	100
3.5	Common Mistakes in Data Visualization Design	102
	Wrong Type of Visualization	102
	Trying to Display Too Much Information	104
	Using Excel Default Settings for Charts	106
	Too Many Attributes	108
	Unnecessary Use of 3D	109
	Summary	111
	Glossary	111
	Problems	112
CHAPTER 4	Purposeful Use of Color	128
4.1	Color and Perception	130
	Attributes of Color: Hue, Saturation, and Luminance	130

	Color Psychology and Color Symbolism	132
	Perceived Color	132
4.2	Color Schemes and Types of Data	135
	Categorical Color Schemes	135
	Sequential Color Schemes	137
	Diverging Color Schemes	139
4.3	Custom Color Using the HSL Color System	141
4.4	Common Mistakes in the Use of Color in Data Visualization	146
	Unnecessary Color	146
	Excessive Color	148
	Insufficient Contrast	151
	Inconsistency Across Related Charts	153
	Neglecting Colorblindness	153
	Not Considering the Mode of Delivery	156
	Summary	156
	Glossary	157
	Problems	157
CHAPTER 5	Visualizing Variability	174
5.1	Creating Distributions from Data	176
	Frequency Distributions for Categorical Data	176
	Relative Frequency and Percent Frequency	179
	Visualizing Distributions of Quantitative Data	181
5.2	Statistical Analysis of Distributions of Quantitative Variables	193
	Measures of Location	193
	Measures of Variability	194
	Box and Whisker Charts	197
5.3	Uncertainty in Sample Statistics	200
	Displaying a Confidence Interval on a Mean	201
	Displaying a Confidence Interval on a Proportion	203
5.4	Uncertainty in Predictive Models	205
	Illustrating Prediction Intervals for a Simple Linear Regression Model	205
	Illustrating Prediction Intervals for a Time Series Model	208
	Summary	211
	Glossary	211
	Problems	213
CHAPTER 6	Exploring Data Visually	226
6.1	Introduction to Exploratory Data Analysis	228
	Espléndido Jugo y Batido, Inc. Example	229
	Organizing Data to Facilitate Exploration	230

6.2	Analyzing Variables One at a Time	234
	Exploring a Categorical Variable	234
	Exploring a Quantitative Variable	237
6.3	Relationships between Variables	242
	Crosstabulation	242
	Association between Two Quantitative Variables	247
6.4	Analysis of Missing Data	256
	Types of Missing Data	256
	Exploring Patterns Associated with Missing Data	258
6.5	Visualizing Time-Series Data	260
	Viewing Data at Different Temporal Frequencies	260
	Highlighting Patterns in Time Series Data	262
	Rearranging Data for Visualization	266
6.6	Visualizing Geospatial Data	269
	Choropleth Maps	269
	Cartograms	272
	Summary	273
	Glossary	274
	Problems	275
CHAPTER 7	Explaining Visually to Influence with Data	284
7.1	Know Your Audience	287
	Audience Member Needs	287
	Audience Member Analytical Comfort Levels	289
7.2	Know Your Message	292
	What Helps the Decision Maker?	293
	Empathizing with Data	294
7.3	Storytelling with Charts	300
	Choosing the Correct Chart to Tell Your Story	300
	Using Preattentive Attributes to Tell Your Story	304
7.4	Bringing It All Together: Storytelling and Presentation Design	306
	Aristotle's Rhetorical Triangle	307
	Freytag's Pyramid	308
	Storyboarding	311
	Summary	313
	Glossary	313
	Problems	314
CHAPTER 8	Data Dashboards	322
8.1	What Is a Data Dashboard?	324
	Principles of Effective Data Dashboards	325
	Applications of Data Dashboards	325

8.2	Data Dashboards Taxonomies	327
	Data Updates	327
	User Interaction	327
	Organizational Function	328
8.3	Data Dashboard Design	328
	Understanding the Purpose of the Data Dashboard	329
	Considering the Needs of the Data Dashboard's Users	329
	Data Dashboard Engineering	330
8.4	Using Excel Tools to Build a Data Dashboard	331
	Espléndido Jugo y Batido, Inc.	331
	Using PivotTables, PivotCharts, and Slicers to Build a Data Dashboard	332
	Linking Slicers to Multiple PivotTables	343
	Protecting a Data Dashboard	346
	Final Review of a Data Dashboard	347
8.5	Common Mistakes in Data Dashboard Design	348
	Summary	349
	Glossary	349
	Problems	350
CHAPTER 9	Telling the Truth with Data Visualization	360
9.1	Missing Data and Data Errors	363
	Identifying Missing Data	363
	Identifying Data Errors	366
9.2	Biased Data	369
	Selection Bias	369
	Survivor Bias	372
9.3	Adjusting for Inflation	374
9.4	Deceptive Design	377
	Design of Chart Axes	377
	Dual-Axis Charts	381
	Data Selection and Temporal Frequency	382
	Issues Related to Geographic Maps	386
	Summary	388
	Glossary	389
	Problems	389
REFERENCES		397
INDEX		399