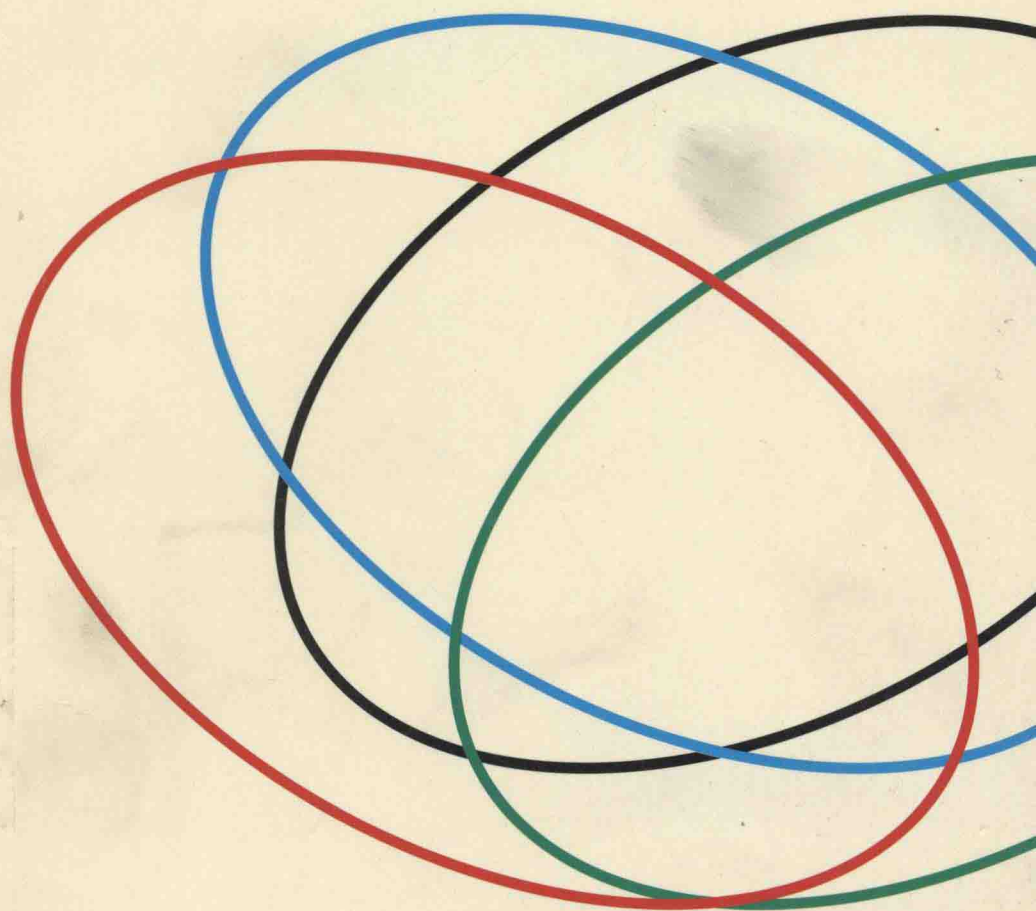


Per Mollerup

# DATA DESIGN

Visualising  
quantities, locations, connections



B L O O M S B U R Y



**Bloomsbury Visual Arts**

An imprint of Bloomsbury Publishing Plc

Imprint previously known as A&C Black Visual Arts

50 Bedford Square  
London  
WC1B 3DP  
UK

1385 Broadway  
New York  
NY 10018  
USA

**[www.bloomsbury.com](http://www.bloomsbury.com)**

**BLOOMSBURY and the Diana logo are trademarks of Bloomsbury Publishing Plc**

© Per Mollerup, 2015

Per Mollerup has asserted his right under the Copyright, Designs and Patents Act, 1988, to be identified as Author of this work.

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage or retrieval system, without prior permission in writing from the publishers.

No responsibility for loss caused to any individual or organization acting on or refraining from action as a result of the material in this publication can be accepted by Bloomsbury or the author.

**British Library Cataloguing-in-Publication Data**

A catalogue record for this book is available from the British Library.

ISBN HB: 978-1-4081-9188-0

PB: 978-1-4081-9187-3

**Library of Congress Cataloging-in-Publication Data**

Mollerup, Per

Data Design: Visualising quantities, locations, connections / Per Mollerup p.cm  
Includes bibliographic references and index.

ISBN 978-1-4081-9188-0 (hardcover) ISBN 978-1-4081-9187-3 (pbk.)  
2012045678

Design by Per Mollerup  
Printed and bound in China

## **Data Design**

Visualising quantities, locations, connections

B L O O M S B U R Y  
LONDON • NEW DELHI • NEW YORK • SYDNEY

Contents

Introduction	Organising knowledge, 7
Field of study	Goals, 9
	Benefits, 10
	Information graphics, 14
	Simplicity, 20
	Milestones, 22
Basics	Beware of visual displays, 34
	Research and common sense, 36
	Gestalt principles of perception, 39
	Psychological principles, 44
	Colour, 46
	Notation, 50
	Reading visual displays, 52
	Choice of display, 53
Visualising quantities, 55	Variables, 56
	Pie charts, 60
	Donut charts, 66
	Vertical bar charts, 68
	Histograms, 76
	Stem-and-leaf plots, 77
	Horizontal bar charts, 78
	Divided-bar charts, 82
	Stacked bar charts, 84
	Step charts, 86
	Bubble charts, 88
	Line charts, 90
	Layer charts, 96
	Radar charts, 98
	Scatter plots, 100
	Slopegraphs, 102
	Picture tables, 104
	Heat maps, 108
	Small multiples, 110
	Don'ts, 112

Visualising locations, 115	<ul style="list-style-type: none"> <li>Thematic maps, 116</li> <li>Map scale, 118</li> <li>Map projection, 120</li> <li>The geographic coordinate system, 123</li> <li>Map orientation, 124</li> <li>Choropleth maps, 126</li> <li>Proportional symbol maps, 128</li> <li>Isopleth maps, 130</li> <li>Cartograms, 132</li> <li>Chorochromatic maps, 136</li> <li>Dot distribution maps, 137</li> <li>Transit maps, 138</li> <li>You-are-here maps, 140</li> </ul>
Visualising connections, 143	<ul style="list-style-type: none"> <li>Trees, 144</li> <li>Decision trees, 148</li> <li>Treemaps, 149</li> <li>Mind maps, 152</li> <li>Positioning maps, 154</li> <li>Timelines, 156</li> <li>Flow charts, 158</li> <li>Concept maps, 160</li> <li>Venn diagrams, 162</li> <li>Carroll diagrams, 164</li> <li>Euler diagrams, 166</li> </ul>
End matter	<ul style="list-style-type: none"> <li>Sources, 168</li> <li>Index, 171</li> <li>Acknowledgements, 174</li> <li>Author bio, 175</li> </ul>

**As knowledge increases amongst mankind, and transactions multiply, it becomes more and more desirable to abbreviate and facilitate the modes of conveying information from one person to another, and from one individual to many.**

**William Playfair**

*The Commercial and Political Atlas, 1786, pVII*

## Organising knowledge

*Data Design* deals with data visualisation to organise knowledge about the world into charts, maps, and diagrams. Data visualisation makes raw data useful. This book is not about data decoration.

Data visualisation constitutes a branch of information graphics, which in turn provides visual explanation in fields as diverse as document design, wayshowing, human-machine communication, and interaction design.

Data visualisation reorganises data, presenting it in ways that answer relevant questions by making data easily perceptible and comprehensible. It domesticates raw data to make it clear and usable as a basis for decisions.

The three subject categories in this book – quantities, locations, and connections – are visualised in different, yet similar ways. Together with hybrid forms, the similarity suggests a parallel discussion, although in separate parts.

*Data Design* is written and designed to assist designers, researchers, and writers to express visually what can better be explained and understood by position, shape, size, and colour than by words, letters, and numbers. It describes problems, principles, and solutions for the visual display of information, and it presents didactic examples from the real world.

*Data Design* concentrates on the principles of data visualisation. The choice of computer, software, and type of coding are up to the reader. Such fast-changing aspects of data design will develop into something different from today's versions in the near future. When they do, the principles in *Data Design* will still be valid.

First, *Data Design* discusses the theory behind visual displays and presents basic principles. Second, *Data Design* presents an array of well-known and lesser-known display formats. Together, these principles and practices will enable readers to make useful and engaging visual displays off the rack, along with possible new forms of data design.

Per Mollerup  
Melbourne and Copenhagen, 2014

Most captions in *Data Design* are placed a short distance from the illustrations they explain. When learning about the law of proximity, readers might wonder why. There are two reasons for this apparent failure to obey our own laws. The first reason is purely didactic. Separating illustrations and captions gives readers a chance to see and understand the intentionally self-explanatory charts, maps, and diagrams undisturbed. The second reason is that it adds to the sense of order, ensuring that the displays are not embraced by words.

The text of some of the imported visual displays in *Data Design* is difficult to read. The book format and the resolution of the illustrations don't allow bigger illustrations. While text in displays should normally be perfectly readable, readability of the display text is less important in *Data Design* since it is the principles that matter.

**Goals**

However different in its visual results, data visualisation shares goals with writing. When we write, we do so with one or more of three goals in mind.

First, we write to communicate: to share thoughts, information, and desires with others. In a certain sense, this is the most demanding goal of writing, as it involves the receiver’s perception and cognition. The latter implies that the sender and receiver of communication share a code. For researchers, writing to communicate is important in so far as we require good research to be accessible to others.

Communicate

The second goal of writing is to record what otherwise might be forgotten. This is our goal when we keep a log or diary, produce a list of things to do, write an aide-memoire, or prepare a manuscript for a speech.

Record

The third goal of writing is the most interesting, particularly to researchers. We write to understand. Serious writing is a conversation with the person we know better than anybody else: ourselves. Seeing our thoughts on paper helps us to improve these thoughts and inspires us to think new thoughts. Many of us would not be able to think as well as we do without writing, and without reading what we write.

Understand

Data visualisation serves the same three goals as writing: to communicate, to record, and to understand, with the emphasis on communicating and understanding.

**Benefits**

It takes energy, thought, and time to make good visual displays that can substitute for stacks of naked numbers and lengthy verbal descriptions. These human and economic costs can be justified by at least four benefits:

– *Understandability*

Visual displays present qualitative and quantitative relationships in ways that are easier and faster to understand than text and numbers alone. A written explanation of the connections between London’s 270 Underground stations would take 72,630 (270x269) individual explanations; even more with alternative routes. The iconic Tube map (*see p21*) provides all this information on a single sheet in a pocket-sized format. Most importantly, the map is easily accessible to users.

– *Insight*

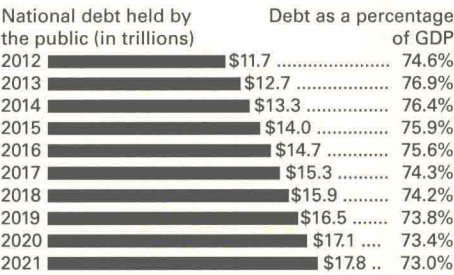
Visual displays reveal relationships and trends that may be difficult to see when the subject matter is explained only in words, letters, and figures.

**Figure 1** Many visual displays designed for fast comparison are supplemented with numbers to offer exact comparison. These displays use the best of both worlds: the speed of visual displays and the exactness of figures. They are hybrids between visual displays and tables. The visual part of this horizontal bar chart shows the growth of US national debt in absolute terms, while the figures also show that the growing debt is decreasing compared with the GDP. Data and design concept sourced from *USA Today*, USA, 21 Sep 2011

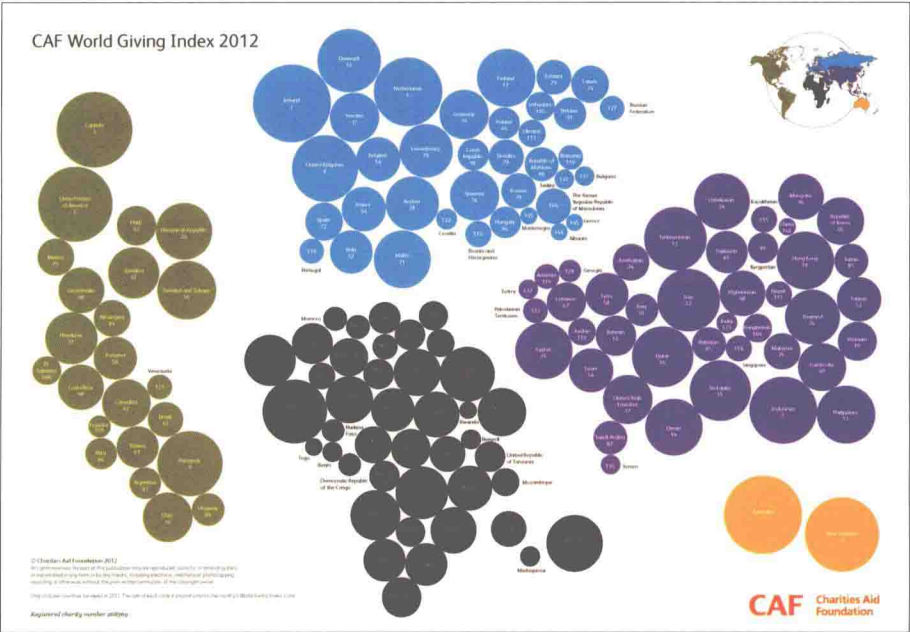
**Figure 2** This Dorling cartogram shows the nations of the world as circles sized according to the charity-giving of the population. The map is able to tell us faster than a table would be able to which nations are the most charitable and how a specific nation ranks. CAF World Giving Index 2012, Charities Aid Foundation. Available from [www.cafonline.org/pdf/WorldGivingIndexA3Map2012WEB.pdf](http://www.cafonline.org/pdf/WorldGivingIndexA3Map2012WEB.pdf)

# Living within means?

Even if recommendations President Obama made Monday are enacted, the national debt would increase \$6 trillion over the next decade.



1



2

## Benefits

### – Attention

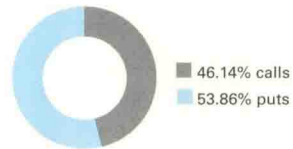
Well-crafted visual displays capture and hold the attention of the audience. Capturing attention is a legitimate reason for data visualisation. In principle, researchers choose visual displays for understandability and insight. Media editors frequently choose visual displays to invigorate their products.

### – Memorability

Simplicity and elegance can make visual displays more memorable than other means of expression. “To remember simplified pictures is better than to forget accurate numbers,” as stated by Otto Neurath, the creator of Isotype (*see p30*).

The first two benefits, understandability and insight, deal with the way the reader understands the subject matter; the last two benefits, attention and memorability, deal with the way the visual displays impact on the reader’s scarce attention and memory. Different situations call for different combinations of these benefits.

**Put/Call-Feeling**



3

**Figure 3** Donut chart as seen in *Neue Zürcher Zeitung*, the well-known Swiss newspaper.

The role of the donut chart on a page filled with text on financial matters is to draw the reader’s attention to the fact that the market is slightly less optimistic (calls) than pessimistic (puts). The figures need no explanation. Data and design concept sourced from *Neue Zürcher Zeitung*, CH, 21 Sep 2011

**To remember simplified pictures  
is better than to forget accurate  
numbers.**

**Otto Neurath**

*Empiricism and Sociology, 1933/1973, p220*

Field of study

**Information graphics**

Data visualisation constitutes a branch of information graphics, a term which is not totally precise. Taken literally, information graphics include all kinds of graphic design, but the commonly used restricted sense of the term designates a specific type of graphic design that works with objectivity and clarity of expression.

– *Definition*

IIID – the International Institute for Information Design – defines information design as

‘The defining of the requirements governing the selecting, rendering, and transmission of information for the purpose of knowledge transfer as well as the optimisation of the information with respect to these requirements.’

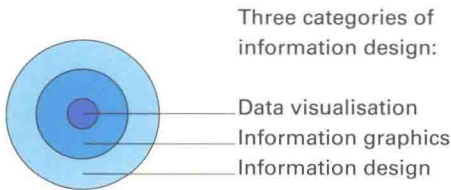
While this definition can, in principle, include disciplines other than visual design, we shall concentrate on graphic design. For clarity, we can define information graphics as the

design of visual explanation

where the term *design* covers the design process as well as the result of that process.

“A definition will often be more comprehensible for being made longer, but any gain in precision tends to be offset by the reader’s resulting muddlement.”

Henry Hitchings  
*Dr Johnson’s Dictionary*,  
John Murray, London, 2005,  
p81



– *Problems addressed*

Information graphics typically deal with issues that are both complex and complicated. Complex means consisting of many interrelated elements. Complicated means difficult to understand or analyse. Complexity often results in complication. Complexity is an objective quality, it deals with facts; complication is a subjective quality, it depends on the reader's perceptive and cognitive capacities. What is complicated to one reader may be simple to another. The word *reader* is here used synonymously with *receiver*.

– *Outcomes*

Information graphics are descriptive, or prescriptive, or both. Descriptive information presents the facts of the world. For example, a line chart showing the development of a nation's export is descriptive. It deals with knowledge and enables sense making. Prescriptive information tells the receiver what to do. For example, user instructions are prescriptive. Prescriptive information deals with action and suggests following rules.

The boundary between description and prescription is not always sharp. Two signs, one saying *Staff only* and the other *Do not enter*, present basically the same information, but are categorised respectively as description and prescription. A single geographic map can offer both a general description of an area and give special route information, telling readers how to get from A to B. The displays presented in *Data Design* are – with very few exceptions – descriptive.

– Objectivity

Information graphics are intentionally objective. The information describes facts not coloured by the feelings of the sender, and addresses the understanding rather than the feelings of the reader. It can be argued that all information graphics are subjective to some degree, since humans design them. However, information designers consciously strive to make information as objective as possible.

Information graphics are concerned with information located in two of the six content fields in the information matrix\* below: objective prescription and objective description. Data visualisation deals primarily with objective description (the result – visual displays – is printed in bold in the matrix).

\*Inspired by  
Pierre Guiraud  
*Semiology*,  
Routledge, London, 1975

Six categories of information – with examples			
Category >	Identification	Prescription	Description
Objective >	ID number	User instructions	<b>Visual displays</b>
Subjective >	Clothing	Advertisement	Painting

– Clarity

Clarity is key to good information graphics, which are intentionally unambiguous, corresponding to what Umberto Eco describes as *closed text*: not open to free interpretation.