

Crash Course in Data Visualization

With Walter McHugh, Head of Data Operations US & Lecturer at Northeastern University



1. What is data visualization?

What is data viz?

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

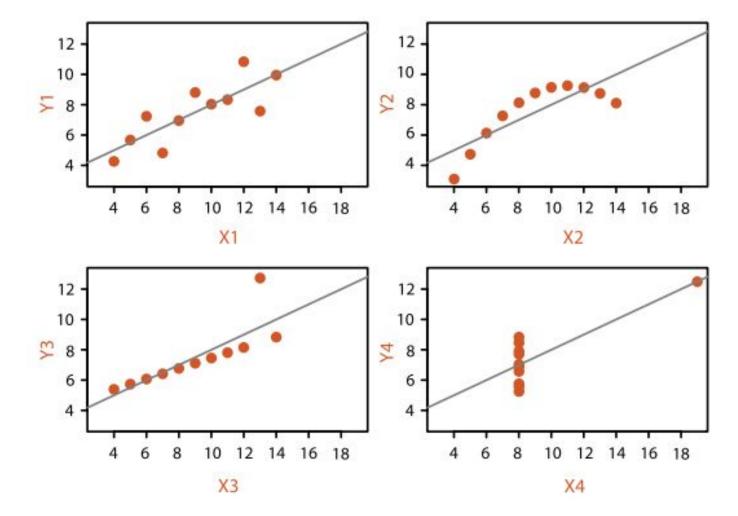
What is data viz?

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

Why do we viz?

Anscombe's Quartet: Raw Data

	1		2		3		4	
	X	Υ	X	Υ	X	Υ	X	Υ
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	



2. How to focus audience attention

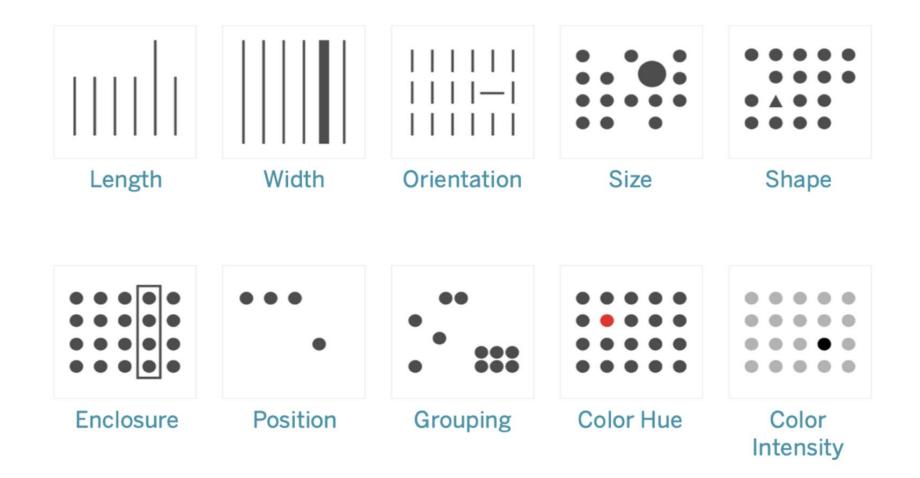
Pre-attentive attributes

How should our audience interact with our visualization?

How can we focus their attention?

Count the 8s

There are eleven



Textual pre-attentive attributes

Bold Color Italics Size

separation

Outline (enclosure)

<u>Underline (added marks)</u>

color, and position

Size

Size, color, and position

State Level A	d Platforn	n Ranking: To	op 3								
		Ad	d Platforr	n							
State	Google	Facebook	Bing	Twitter	Outbrain	State	Google	Facebook	Bing	Twitter	Outbrain
Arizona	1	2	3	4	5	Arizona	1	2	3	4	5
Mass.	1	2	5	4	3	Mass.	1	2	5	4	3
Vermont	2	1	5	4	3	Vermont	2	1	5	4	3
Wisconsin	3	2	5	1	4	Wisconsin	3	2	5	1	4
Washington	1	2	5	4	3	Washington	1	2	5	4	3
Georgia	4	5	4	2	1	Georgia	4	5	4	2	1
Virginia	3	1	2	4	5	Virginia	3	1	2	4	5
Cali.	2	3	4	5	1	Cali.	2	3	4	5	1
Ohio	1	3	5	2	4	Ohio	1	3	5	2	4
lowa	2	3	4	1	5	lowa	2	3	4	1	5

Size, color, and position

Use visual cues, or your audience may look at visuals in a "zig zag" pattern

3. We are all designers

Sullivan: Form follows function

Form: what the user is interacting with

Function: what the user is intended to do

"Form follows function" for dataviz

What do we want our audience to do with our data graphics?

How can we design the graphics to enable this function?

LONDON RAILWAYS REFERENCE METROPOLITAN
ELECTRIC TRAMS,
TOWNOOGREEN,
ALEMANDA PHIACE GOLDERS BAKERLOO RY TO FINCHLEY, TO TOTTENHAM CENTRAL LONDON RY ABARNET. CITY & SOUTH LONDON RY. DISTRICT RAILWAY. GREAT NORTHERN & CITY RY. FINSBURY HIGHGATE PARK HAMPSTEAD RY METROPOLITAN RY. MAMPSTEAD PICCADILLY RY GILLESPIE LONDON UNITED TRAMS. TUFNELL OTHER TRAMWAYS. SUBWAY CONNECTIONS, DRAYTON PARK OTHER RAILWAYS. BELSIZE HOLLOWAY KENTISH HIGHBURY FINCHLEY RO & CALEDONIAN KINGSBURY CHALK SOUTH KENTISH TOWN MARROW, UXORIDGE, SWISS MARROW, UNGRIDO PINNER, MICHMAN SWONTH, CHORLEY WOOD. CHALFONTHO, CHESNAH, AVIESBURY, ETC. COTTAGE YORK ESSEX RO CAMDEN ROAD MARLBORO' ROAD KINGS CROSS MORNINGTON PANCRAS ANGEL STJOHN'S WOOD RO EUSTON GOWER ST OLD ST MOORGATE WHITECHAPEL ALDERS GATE WARREN RUSSELL SQUARE ROYAL OAK GREAT CENTRAL BISHOPS EDGWARE FARRINGDON HOORGATE GOODGES POST BAKER ST WESTBOURNE BRITISH BANK ALDGATE HOLBORN MERY LANE ALDGATE BONDS PRAEDS STRAND EXHIBITION NNONS? MARK LANE WOOD LANE ACTON EALING SOUTHALLUX BRIDGE ETC OUEENS RO DOVERST LONDON BRIDGE NOTTING HILL SAS DOWN ST SHEPHERDS DESTRICT AT TO HYDE PARK CORNER BUSH ACTON, EALING, MOUNSLOW, MARROW, ETC CHARING CROSS KNIGHTSBRIDGE BOROUGH KENSINGTON PATERICO TURNHAM RAVENSCOURT WESTMINSTER GREEN BROMPTON STJAMES PARK STEWS RICHARDS VICTORIA ARMY & NAVY UNITED TRAMS KENSINGTON TWICHENHAM HOUNSLOW SQUARE KENNINGTON MEW.ETC WEST WALHAM GREEN

PARSONS

STOCKWELL FOR BRIXTON

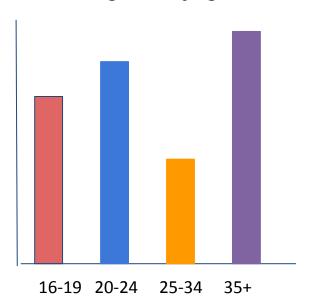
JOHNSON RIDDLE & C+ LTP LONDON'S E

CLAPHAM RE

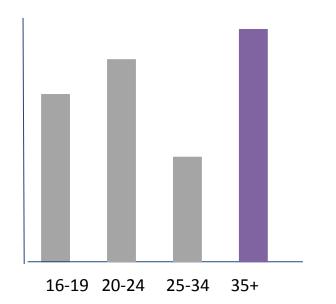


De-emphasize what's not important

Click-through rate by age bracket

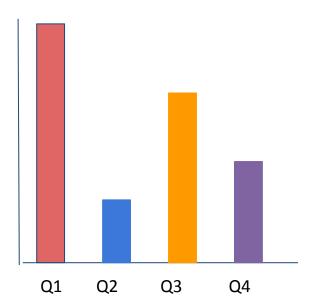


35+ has the highest click-through rate

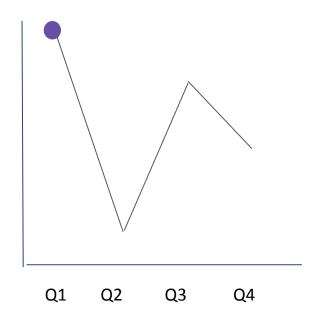


Clean it up

Click-through rate for 35+ by quarter



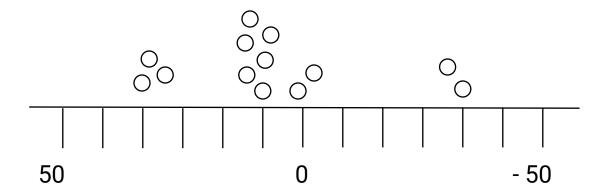
Q1 has highest click-through rate



Avoid overcomplicating

- Use legible fonts consistently
- Strive for "clean" aesthetic
- Use simple language
- Avoid complexity

Consider your audience's intuition



4. What is a dashboard?

What is a dashboard?

"A dashboard is a visual display of the most important information needed to achieve one or more objectives that has been consolidated in a single computer screen so it can be monitored at a glance."

- Stephen Few, Information Dashboard Design (2006)

What is a dashboard?

- Display information for specific objectives
- Fits on one "screen"
- Frequency of data relates to objectives
- User can monitor information "at a glance"

3 main types of dashboards?

- Strategic for managers and executives; simple visual elements
- Analytical enable deeper data analysis and drill downs
- Operational built for constant change; focus on the "now"