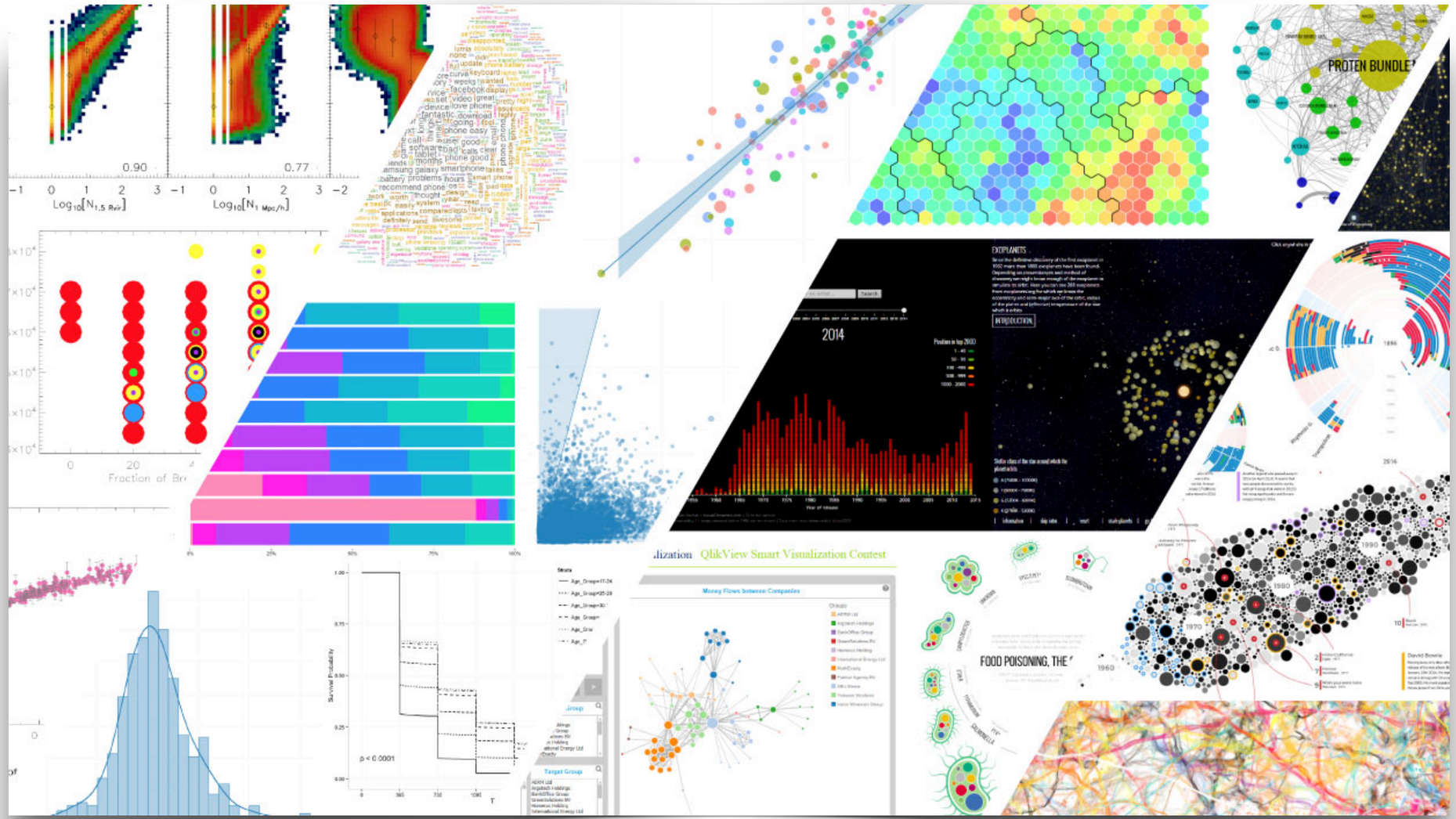


data representation

reading a graph

dataviz *perception* consists in the decoding of the elements of a graph:

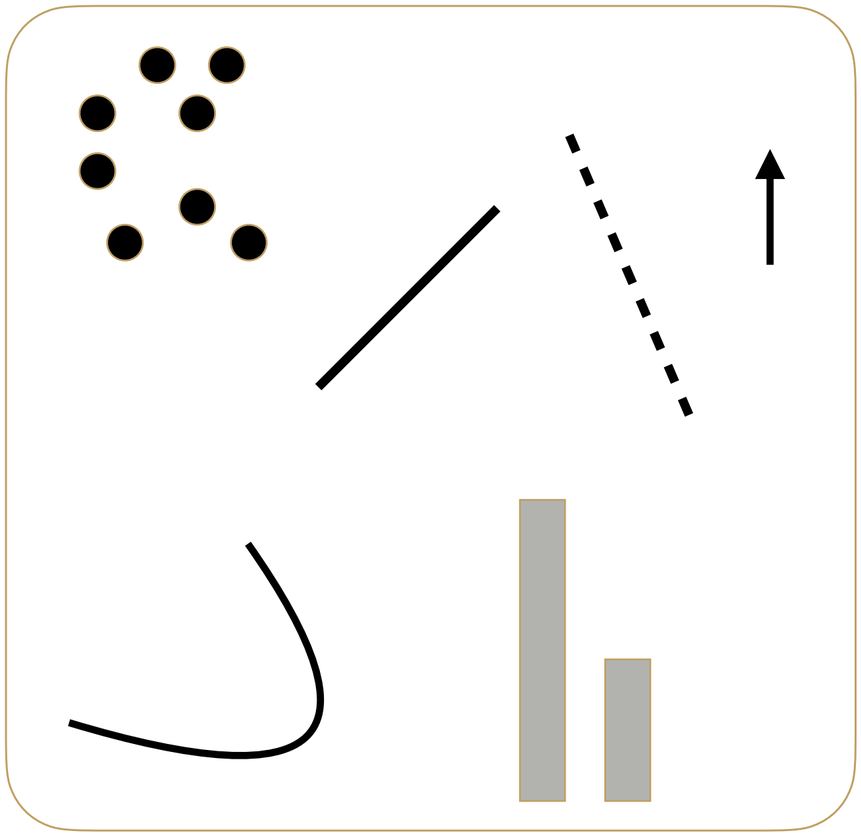
- shapes
- sizes
- positions
- colours

dataviz *interpretation* consists in the decoding of the meaning of a graph, i.e. making sense of the overall graphic construction.

chart = visual encoding + chart apparatus

visual encoding

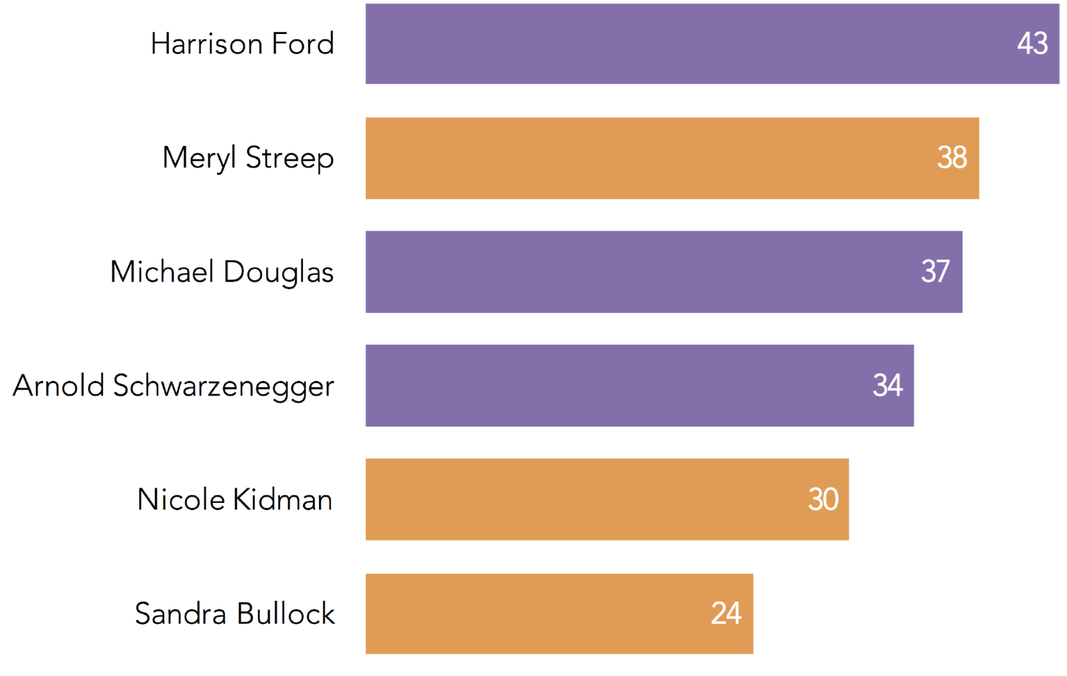
marks



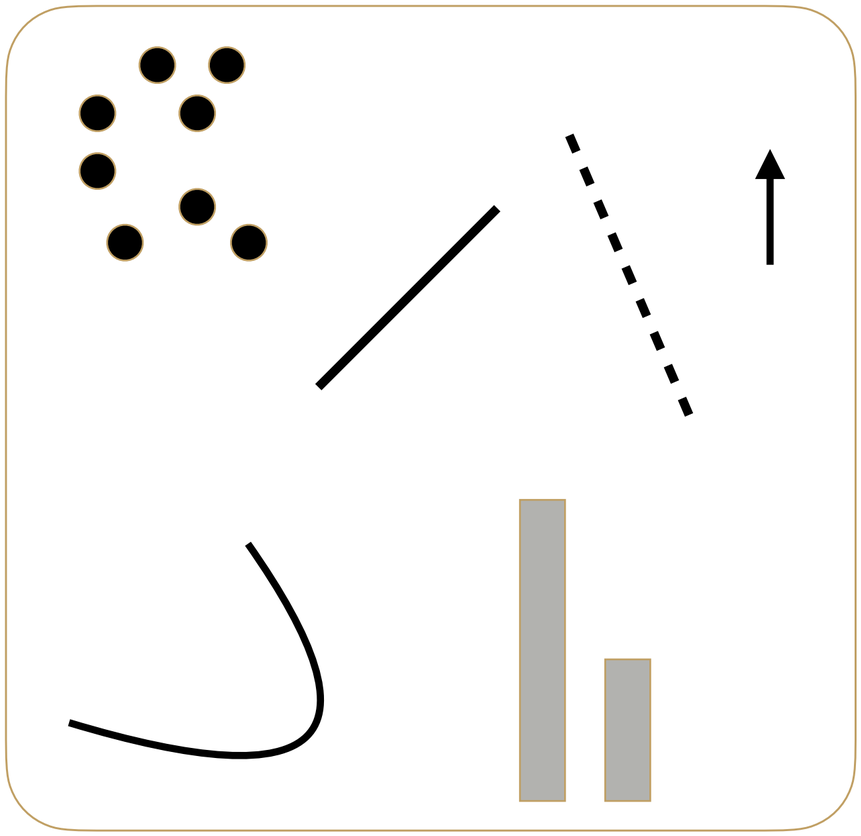
attributes

- quantity
- size
- position
- colour

ACTOR	GENDER	YEARS SINCE FIRST MOVIE
Harrison Ford	Male	43
Meryl Streep	Female	38
Michael Douglas	Male	37
Arnold Schwarzenegger	Male	34
Nicole Kidman	Female	30
Sandra Bullock	Female	24



marks



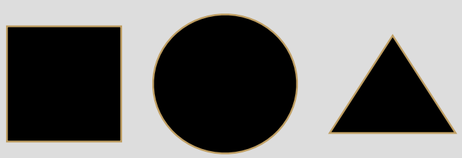
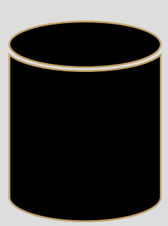
visual encoding

attributes

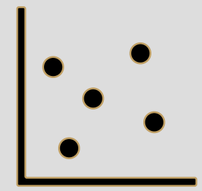
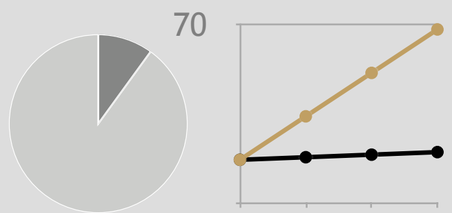
- quantity
- size
- position
- colour

finding the right blend of marks & attributes that most effectively portray the angle of analysis you wish to show

marks

mark	example	description
point		• no spatial variations • quantity as position on a scale
line		• linear spatial dimension • quantity as variation in size
area		• quadratic spatial dimension • quantity as variation in size & position
form		• cubic spatial dimension • quantity as variation in volume

attributes
quantitative 1/2

attribute	example	description
position		• quantity as position on a scale
size		• quantity as variation in size
angle / slope		• quantity as variation of angle • quantity as different slope
quantity		• quantity as repeated set of point marks

attributes
quantitative 2/2

attribute	example	description
colour saturation		• quantity as saturation
colour lightness		• quantity as brightness
pattern		• quantity as density / shape of pattern
motion		• movement as binary indicator to draw focus or to represent quantitative scale ramp

attributes

categorical / relational

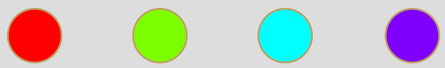
attribute	example	description
symbol / shape		• symbols as categorical association
colour hue		• quantity as brightness
connection / edge		• relationship as connection
containment		• grouping relationship as containment

chart types

bottom-up: from visual encoding to chart type

top-down: choosing a chart type first

a whole world other than the classic 3: bar/pie/line

the chrts classification

categorical	comparing categories and distributions of quantitative values
hierarchical	charting part-to-whole relationships & hierarchies
relational	graphing relationships to explore correlations & connections
temporal	showing trends & activities over time
spatial	mapping spatial patterns through overlays & distortions

caveat: small multiples not included - they are not a different chart on their own, rather an editorial thinking solution

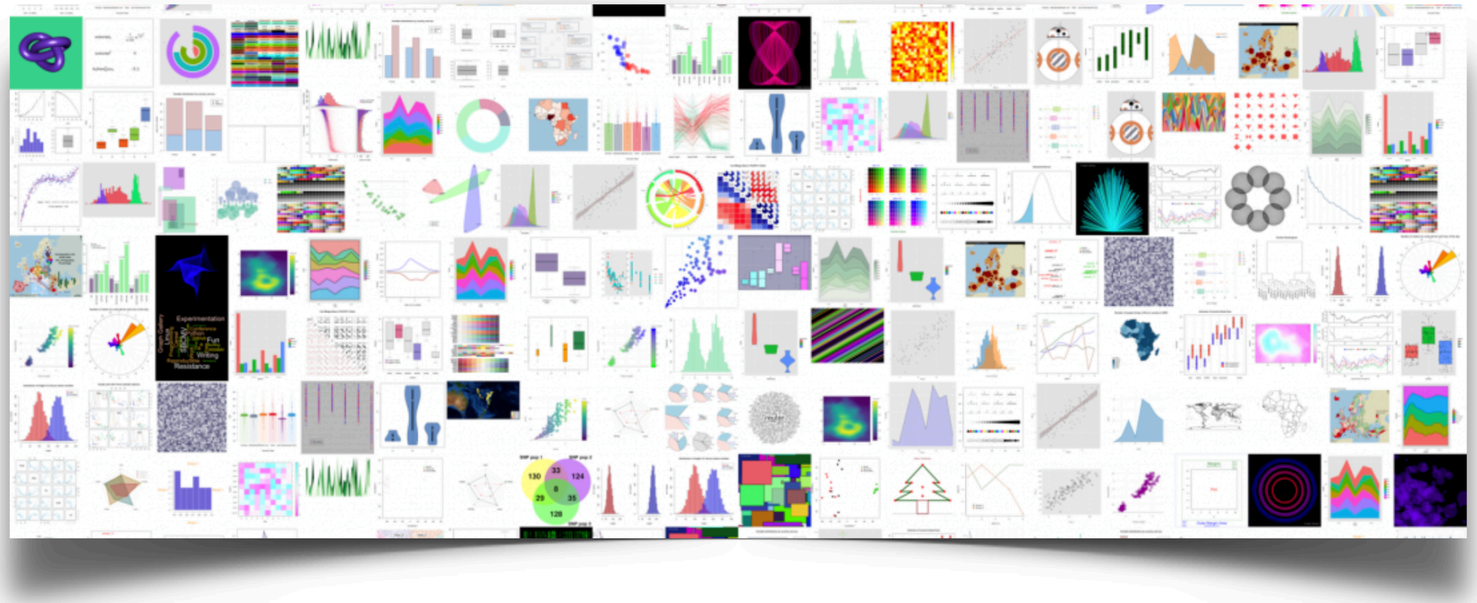


chart zoo

Stage 4

Developing your design **solution**

influencing factors

Stage 1

Formulating your **brief**

Stage 2

Working with **data**

Stage 3

Establishing your **editorial** thinking

Stage 1

Formulating your **brief**

the brief

skill & resources + frequency

which graph can you actually make and how efficiently?

expressiveness

- maximum expressiveness — can create **any** combination of mark & attribute encoding
- limited expressiveness — limited scope, need for workaround, but quick&simple charting

Stage 1

Formulating your **brief**

the brief

purpose

- should you represent your data in a chart form?
- will it add any value, new insights, greater perceptual efficiency w.r.t to non-visualised form?
- will portraying your data in an elegant table actually offer a more suitable solution?
- maybe an information-based solution (imagery, text, video, photo) would work better?

Stage 1

Formulating
your **brief**

the brief

purpose map

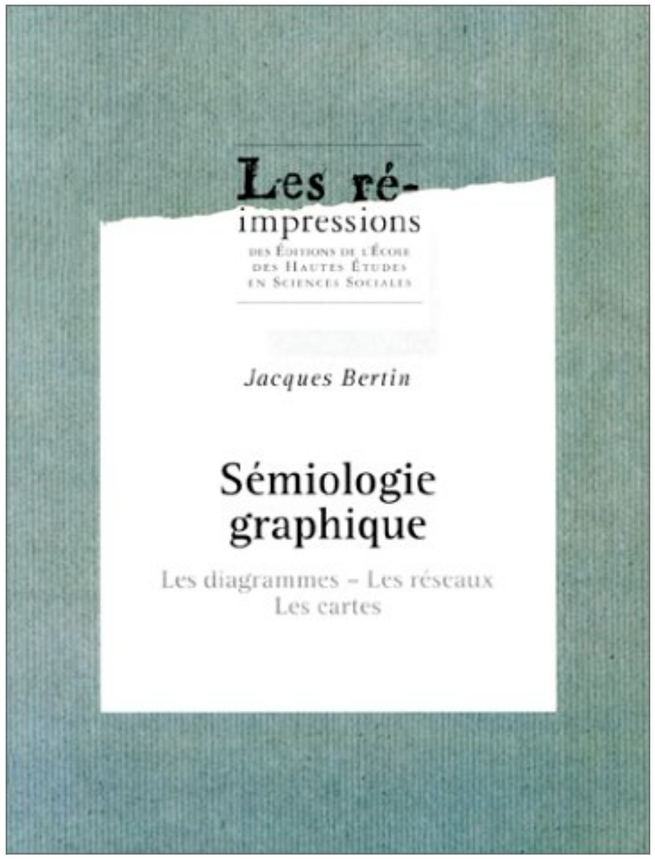
- *reading* of the data or *feeling* of the data?
- precise and accurate perception of values
  or sense-making of the gist of values
- trade-off between emotional qualities and perceptual
  efficiency?

Stage 1

Formulating your **brief**

the brief

purpose map



different ways of encoding data might offer varying degrees of effectiveness in perception
[j. bertin, 1967]

Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods

William S. Cleveland; Robert McGill

Journal of the American Statistical Association, Vol. 79, No. 387 (Sep., 1984), 531-554.

general ranking explaining which attributes would facilitate the highest degree of perceptual accuracy
[w.s. cleveland & r. mcgill, 1967]

Stage 1

Formulating your **brief**

the brief

purpose map

Automating the Design of Graphical Presentations of Relational Information

JOCK MACKINLAY
Stanford University

ACM Transactions on Graphics, Vol. 5, No. 2, April 1986, Pages 110–141.

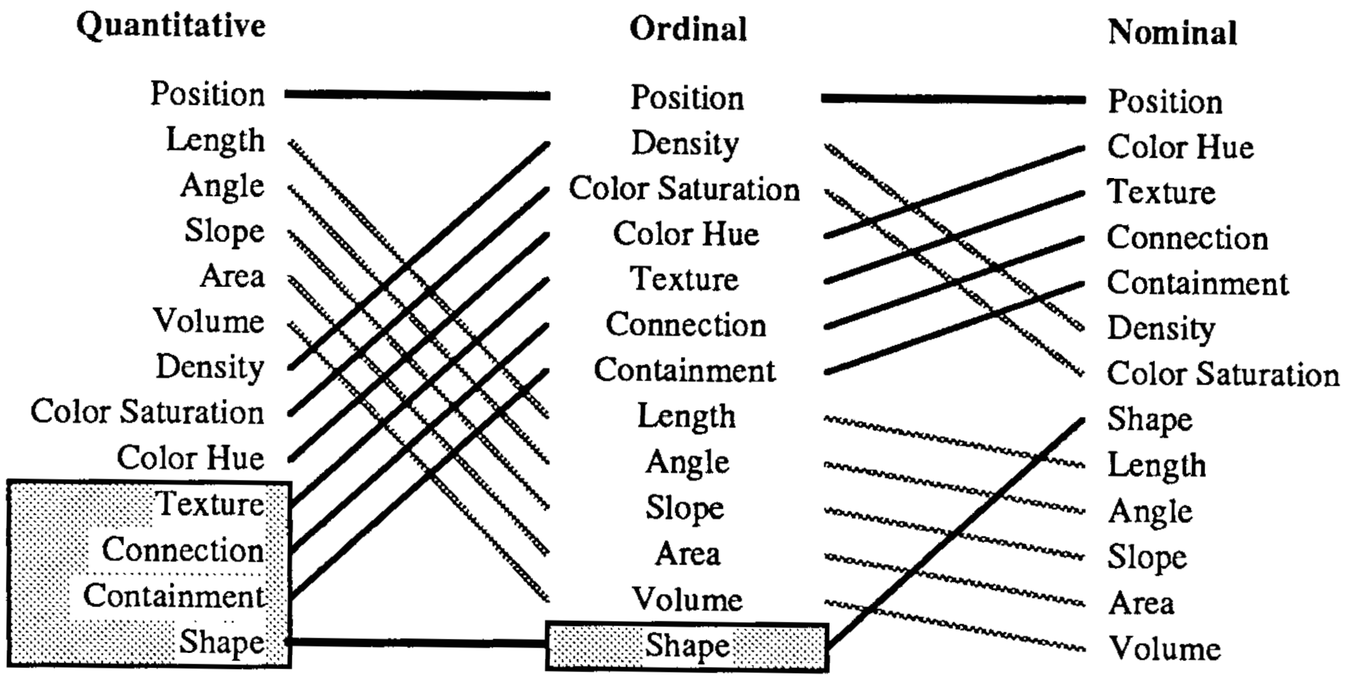
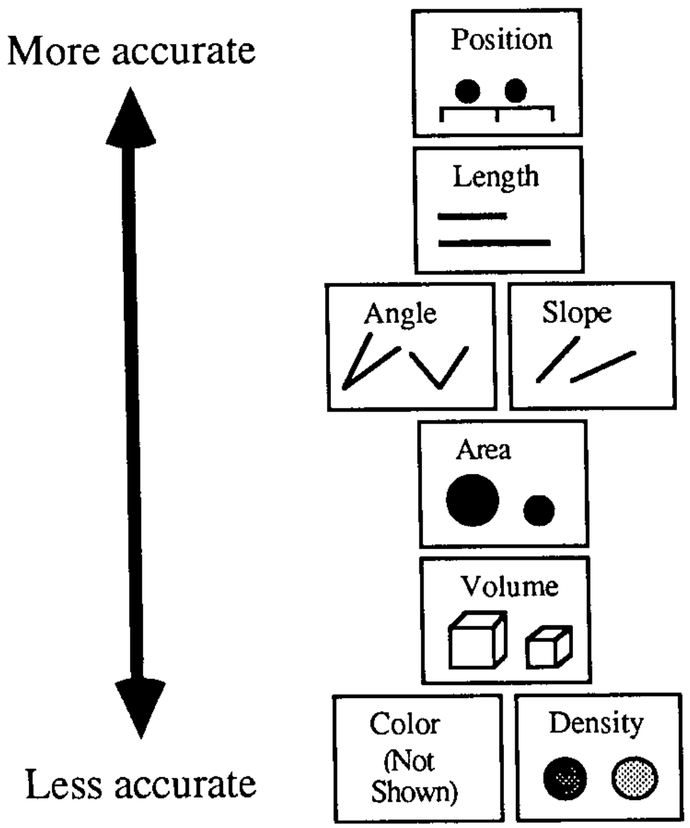


Fig. 15. Ranking of perceptual tasks. The tasks shown in the gray boxes are not relevant to these types of data.



	Nominal	Ordinal	Quantitative
Size	–	•	•
Saturation	–	•	•
Texture	•	•	
Color	•	*	
Orientation	•		
Shape	•		

Fig. 25. Expressiveness of retinal techniques. The – indi-

Stage 1

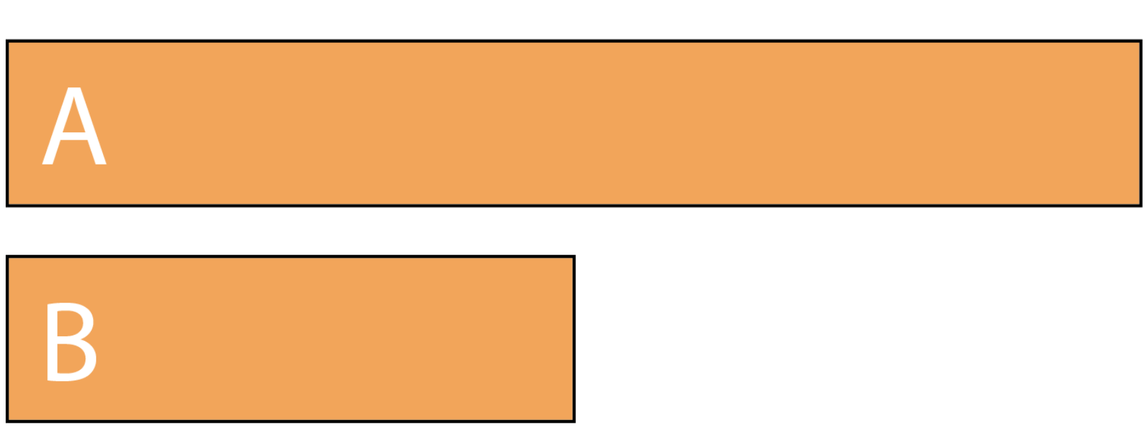
Formulating your **brief**

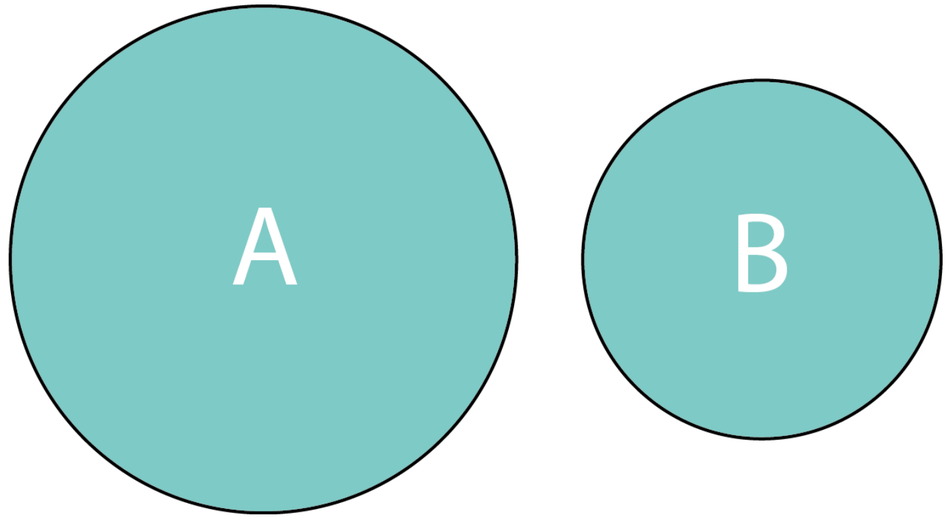
the brief

purpose map

summary: some attributes make it easier and others make it harder to judge accurately the values being portrayed

what's the ratio B/A?





judging variations in lines is far more precise than in areas, even worse in circles: this is due to the 2D measure and the shape
(and in the previous table, *length* is ranked higher than *area*)

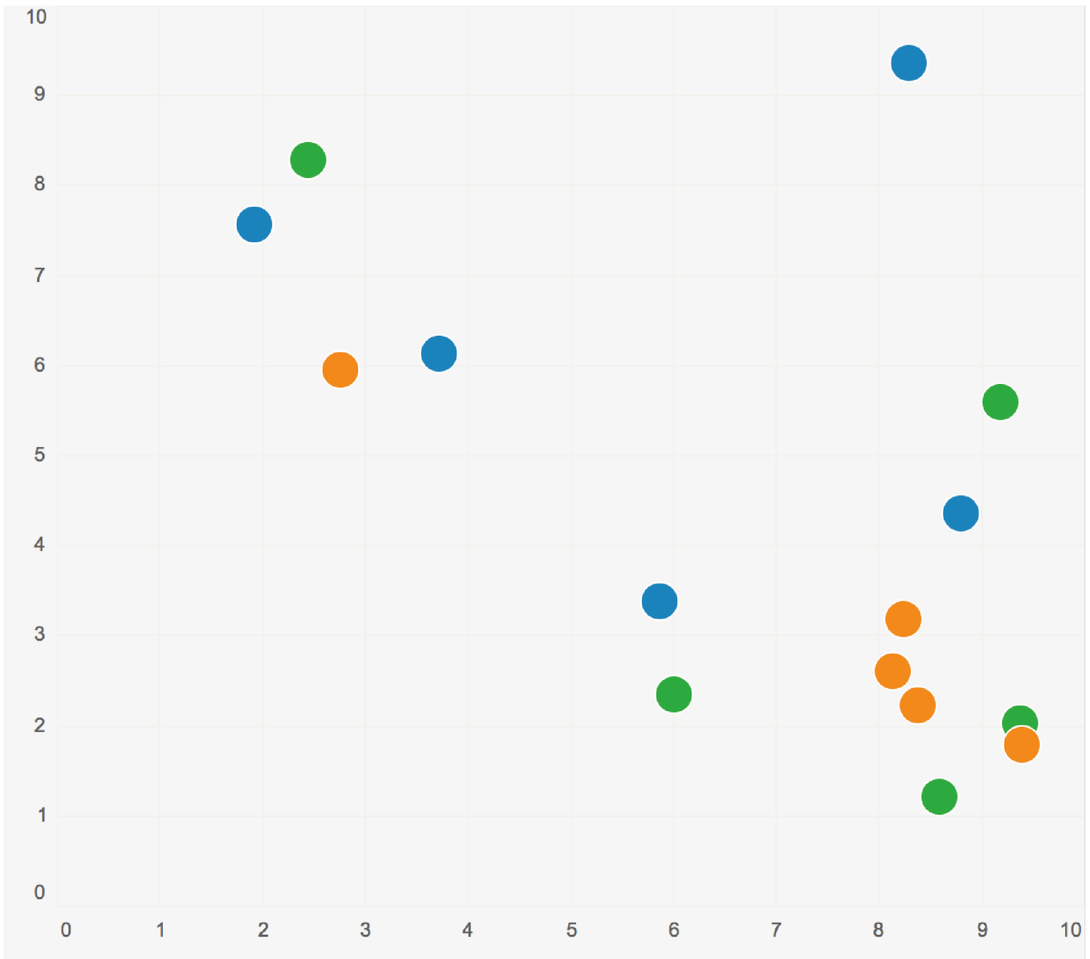
Stage 1

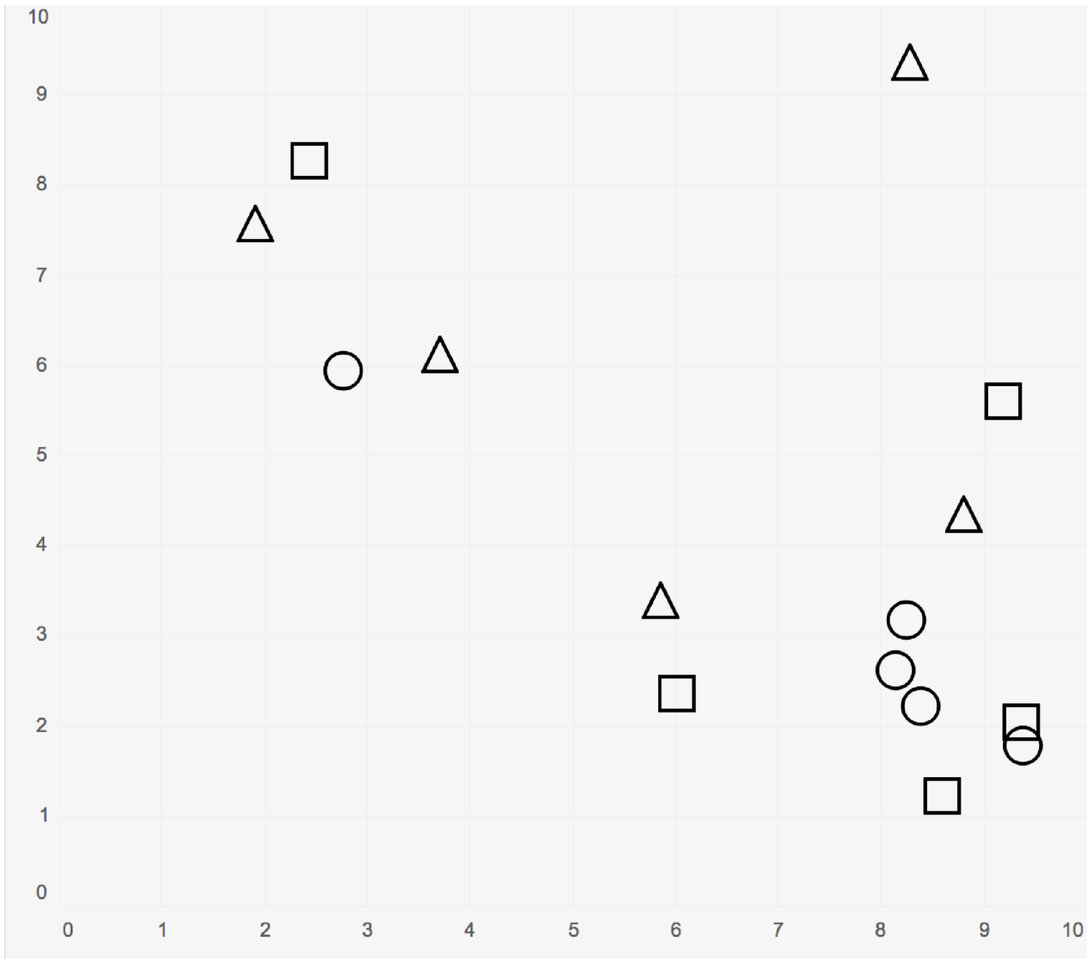
Formulating your **brief**

the brief

purpose map

hue vs *shape* in nominal data





Stage 1

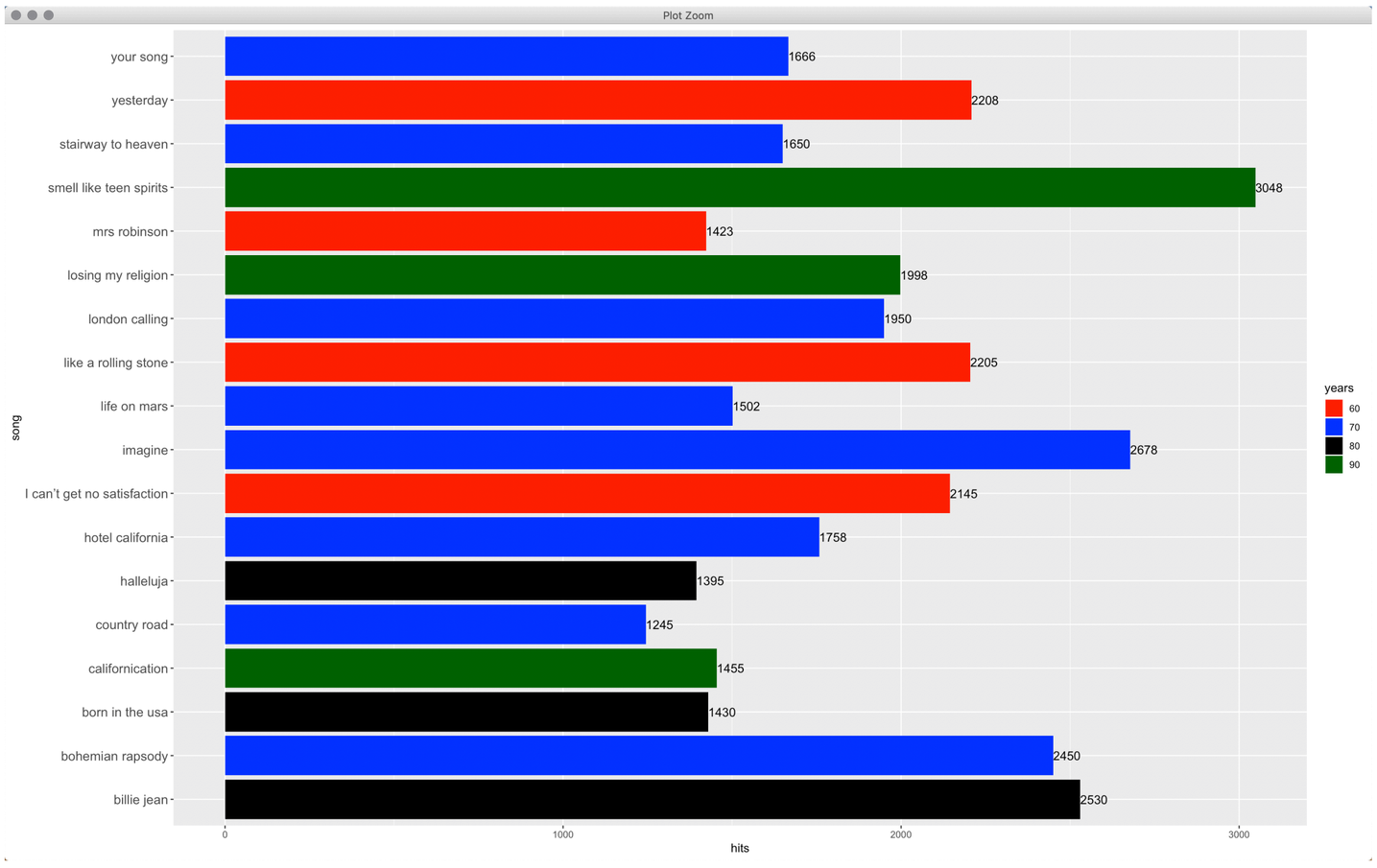
Formulating your **brief**

the brief

purpose map

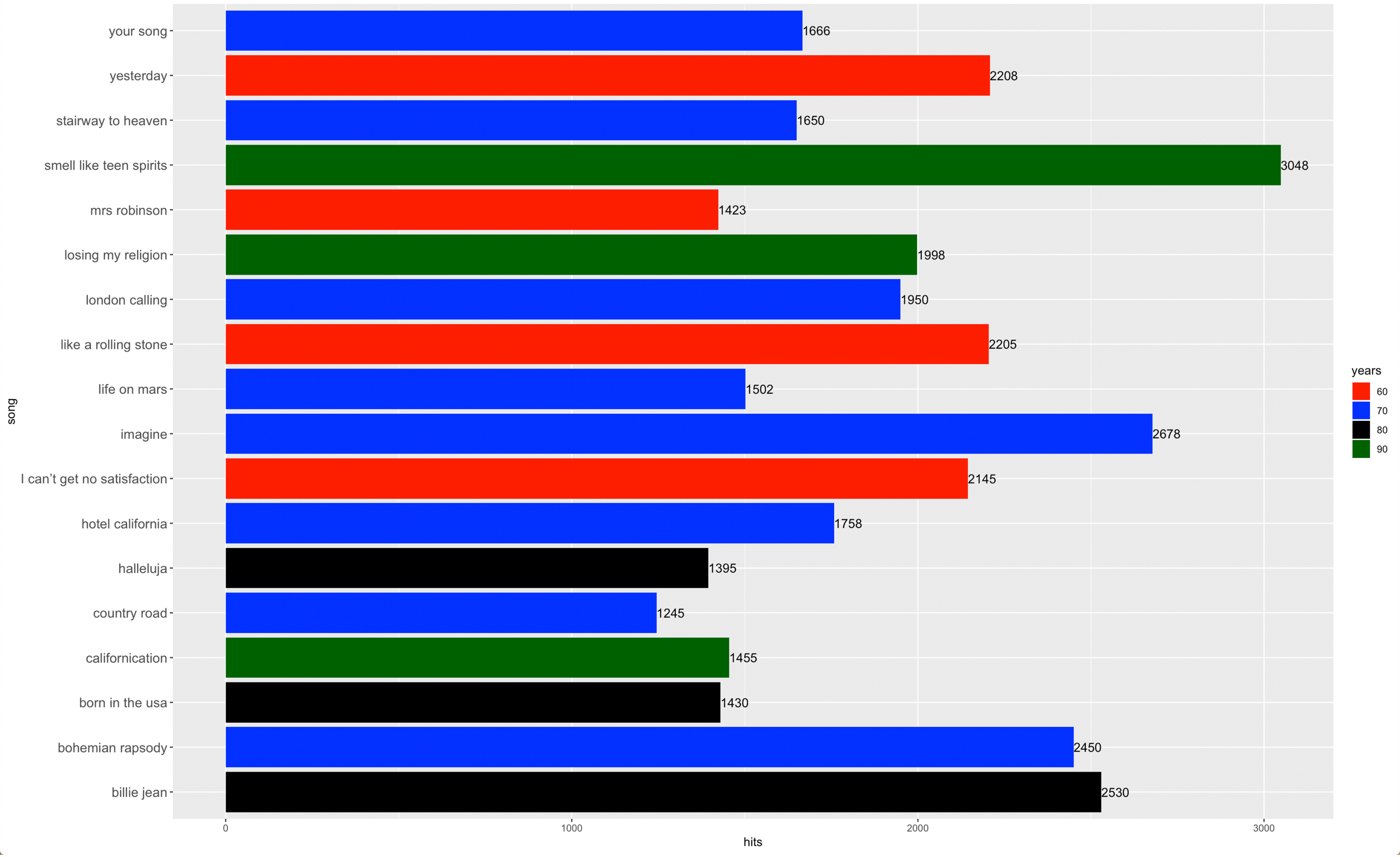
music website poll

precise perception is the aim?



song	yr	hits
smell like teen spirits	90	3048
imagine	70	2678
billie jean	80	2530
bohemian rapsody	70	2450
yesterday	60	2208
like a rolling stone	60	2205
I can't get no satisfaction	60	2145
losing my religion	90	1998
london calling	70	1950
hotel california	70	1758
your song	70	1666
stairway to heaven	70	1650
life on mars	70	1502
californication	90	1455
born in the usa	80	1430
mrs robinson	60	1423
halleluja	80	1395
country road	70	1245

Plot Zoom


your song
yesterday
stairway to heaven
smell like teen spirits
mrs robinson
losing my religion
london calling
like a rolling stone
life on mars
imagine
I can’t get no satisfaction
hotel california
halleluja
country road
californication
born in the usa
bohemian rapsody
billie jean
1666
2208
1650
3048
1423
1998
1950
2205
1502
2678
2145
1758
1395
1245
1455
1430
2450
2530
song
0
1000
2000
3000
hits
years
60
70
80
90

Stage 1

Formulating your **brief**

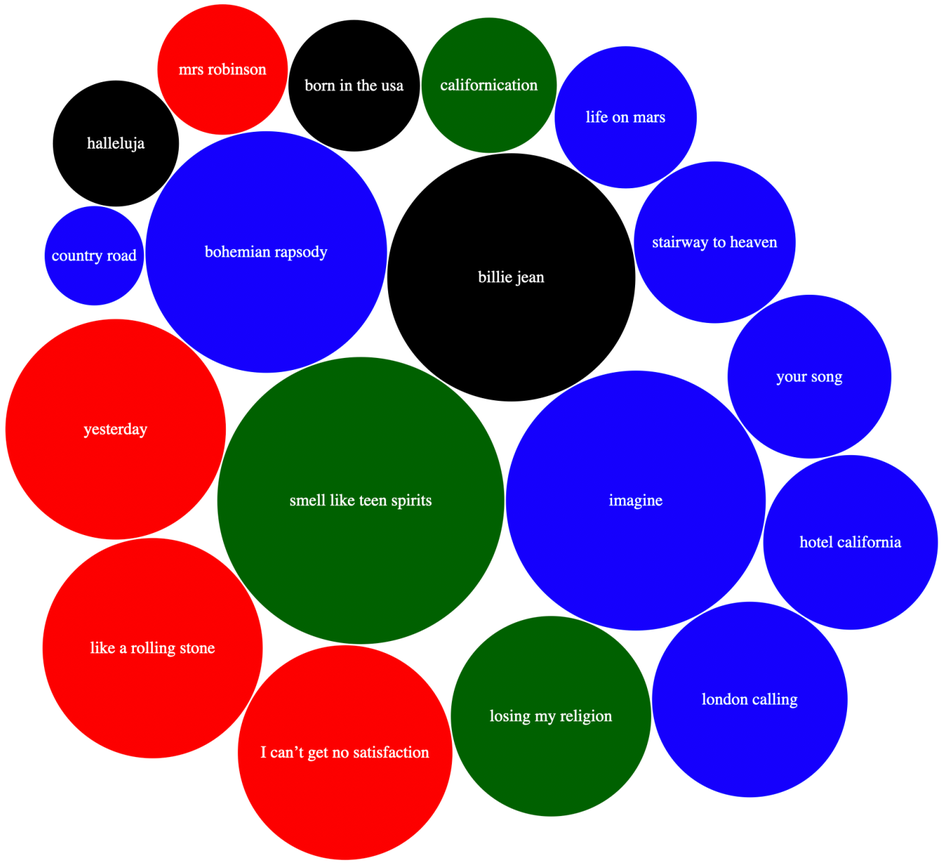
the brief

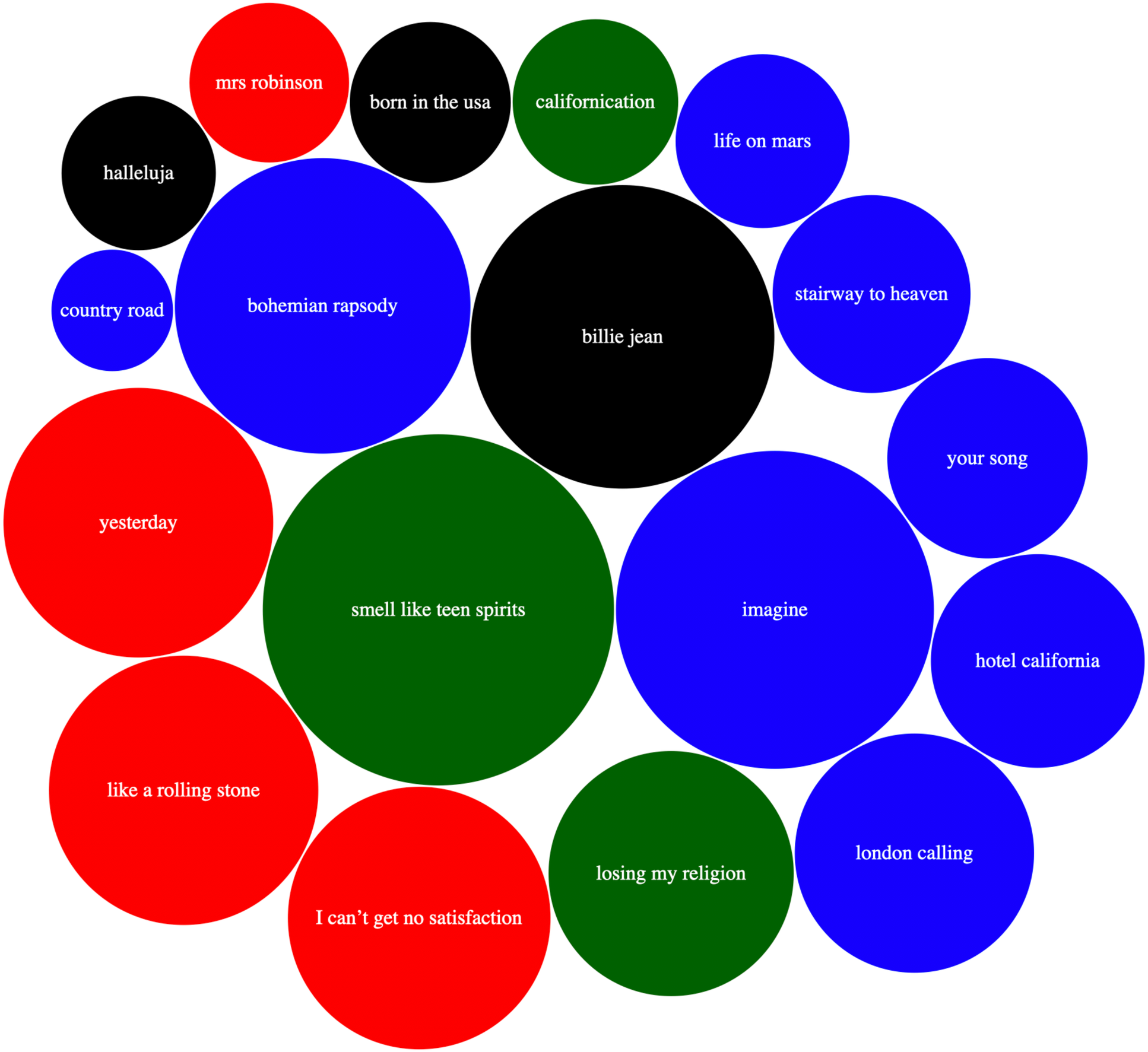
purpose map

music website poll

song	yr	hits
smell like teen spirits	90	3048
imagine	70	2678
billie jean	80	2530
bohemian rapsody	70	2450
yesterday	60	2208
like a rolling stone	60	2205
I can’t get no satisfaction	60	2145
losing my religion	90	1998
london calling	70	1950
hotel california	70	1758
your song	70	1666
stairway to heaven	70	1650
life on mars	70	1502
californication	90	1455
born in the usa	80	1430
mrs robinson	60	1423
halleluja	80	1395
country road	70	1245

or aesthetically give the sense of data?




mrs robinson
born in the usa
californication
life on mars
halleluja
bohemian rapsody
billie jean
stairway to heaven
country road
your song
yesterday
smell like teen spirits
imagine
hotel california
like a rolling stone
I can’t get no satisfaction
losing my religion
london calling

Stage 1

Formulating your **brief**

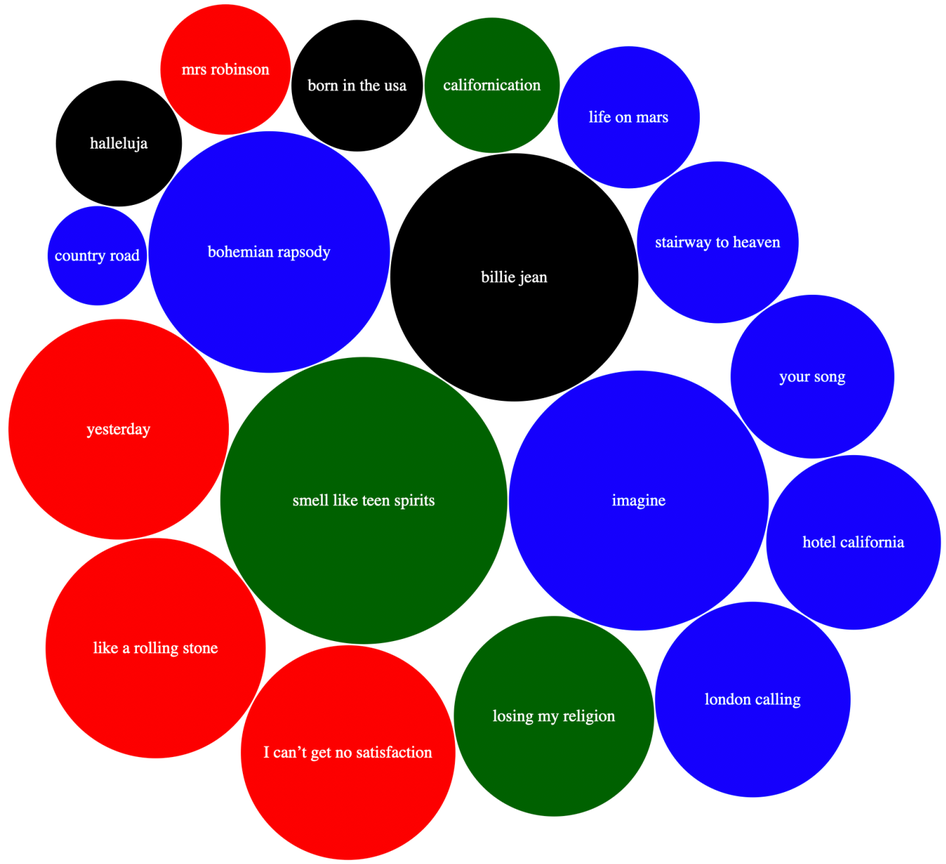
the brief

purpose map

music website poll

song	yr	hits
smell like teen spirits	90	3048
imagine	70	2678
billie jean	80	2530
bohemian rapsody	70	2450
yesterday	60	2208
like a rolling stone	60	2205
I can’t get no satisfaction	60	2145
losing my religion	90	1998
london calling	70	1950
hotel california	70	1758
your song	70	1666
stairway to heaven	70	1650
life on mars	70	1502
californication	90	1455
born in the usa	80	1430
mrs robinson	60	1423
halleluja	80	1395
country road	70	1245

or aesthetically give the sense of data?



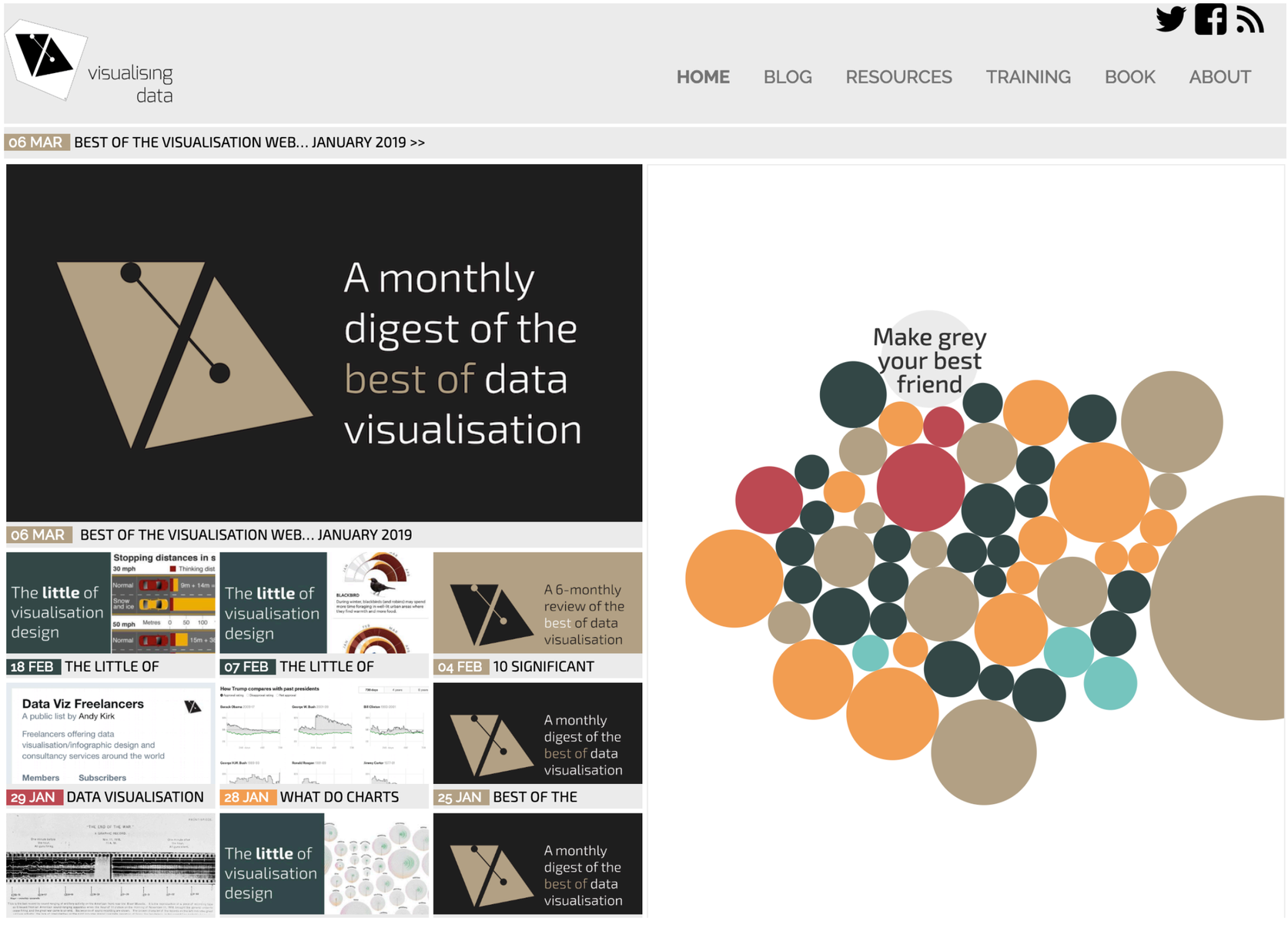
no legend, no figures, bad proportion, transformed data

Stage 1

Formulating your **brief**

the brief

purpose map



andy kirk's visualisation blog

Stage 2

Working with **data**

data processing

only certain types of data can fit into certain chart types, and vice versa

inherent meaning: flower & blossoming metaphor conveys idea of better life

Stage 2

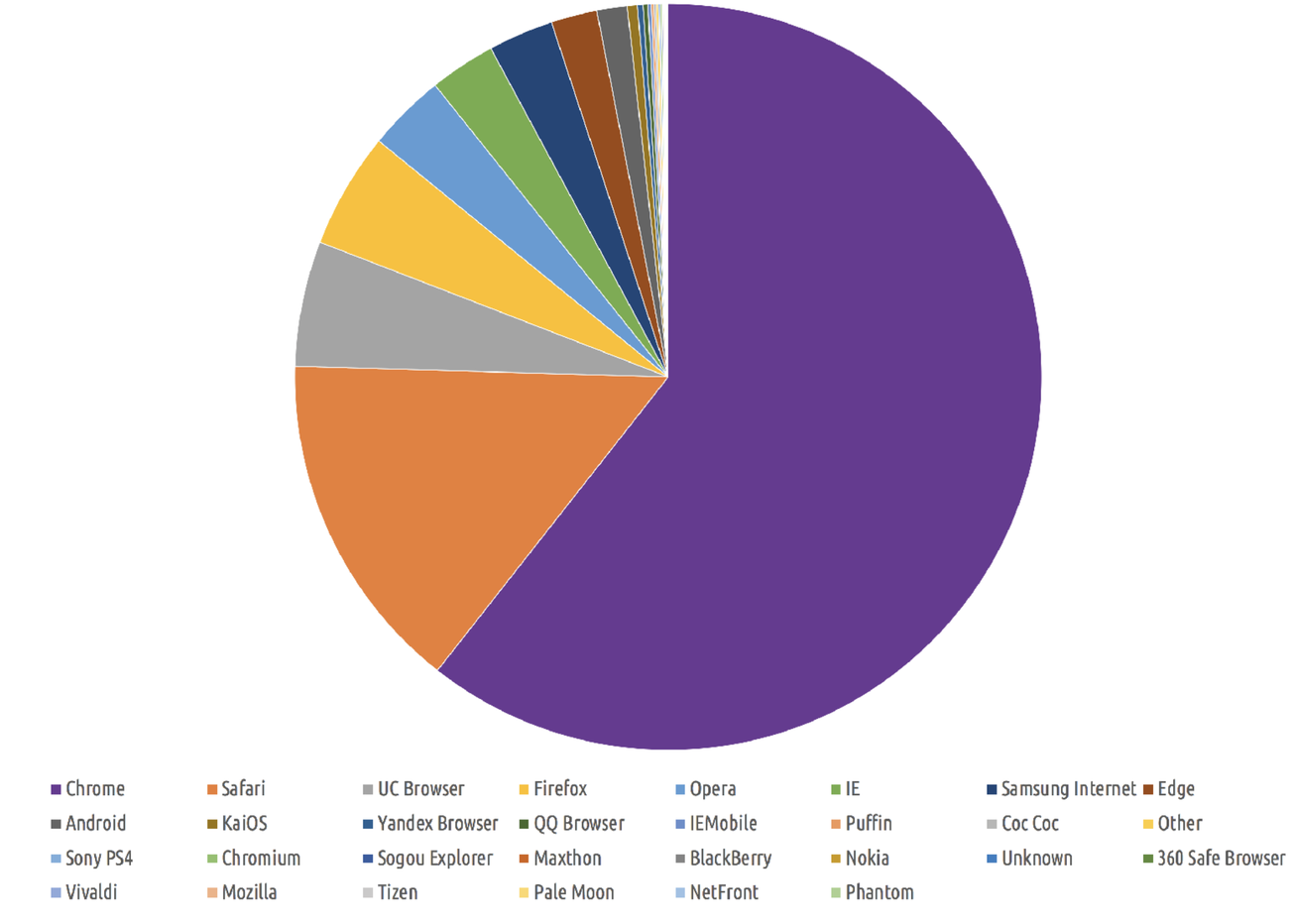
Working with **data**

data processing

only certain types of data can fit into certain chart types, and vice versa

A Chrome Dominates a Cluttered Browser Market

At September 2018, data source: gs.statcounter.com/browser-market-share



Stage 2

Working with **data**

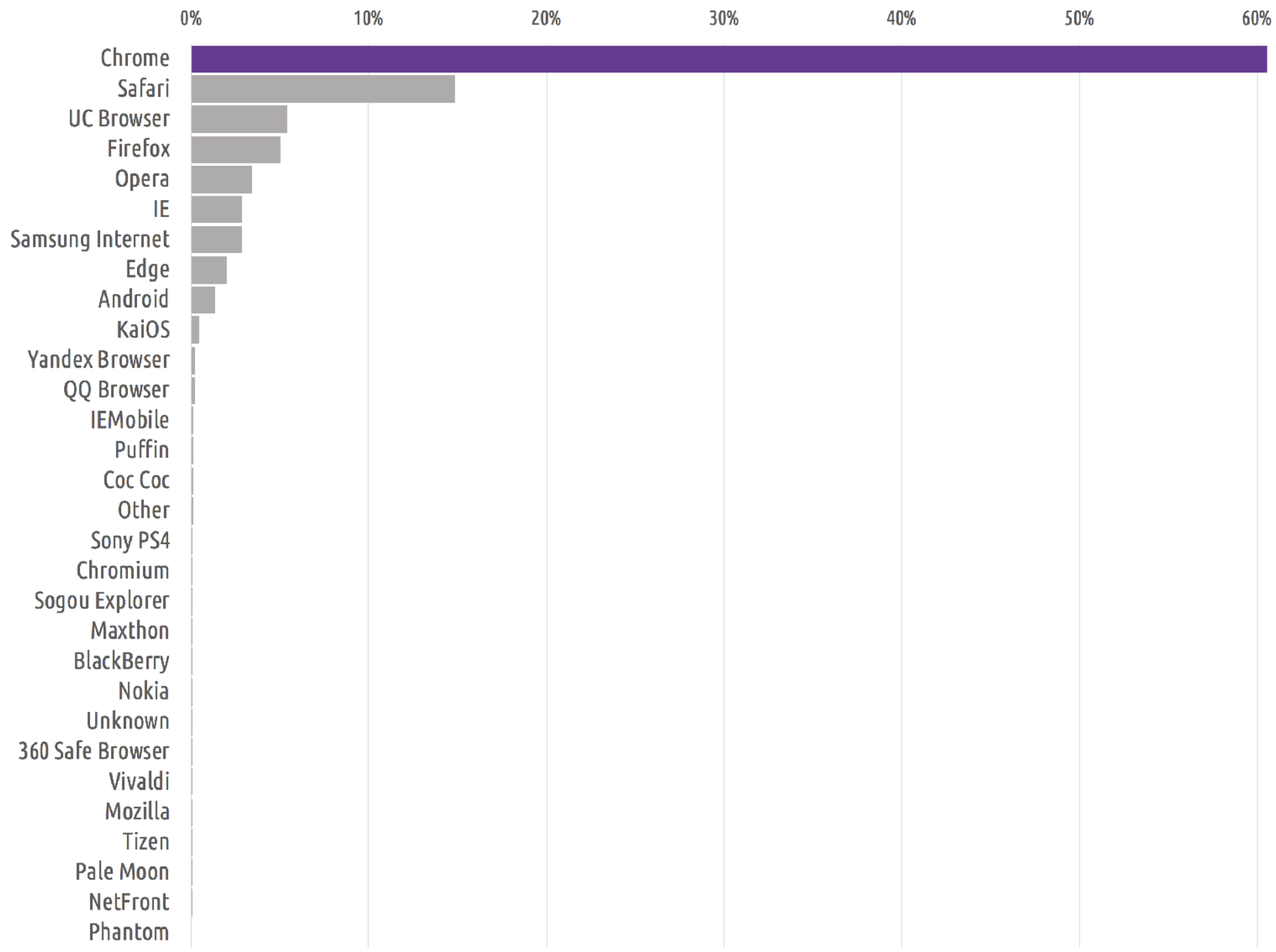
data processing

only certain types of data can fit into certain chart types, and vice versa

B

Chrome Dominates a Cluttered Browser Market

At September 2018, data source: gs.statcounter.com/browser-market-share



Stage 2

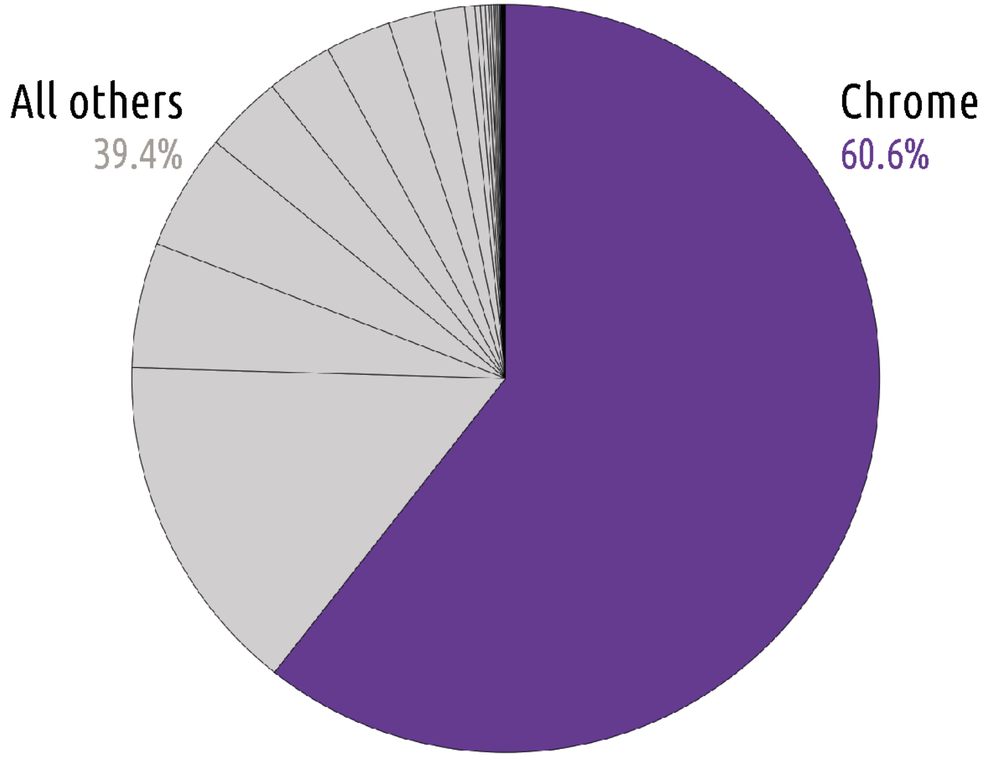
Working with **data**

data processing

only certain types of data can fit into certain chart types, and vice versa

C Chrome Dominates a Cluttered Browser Market

At September 2018, data source: gs.statcounter.com/browser-market-share



Stage 2

Working with **data**

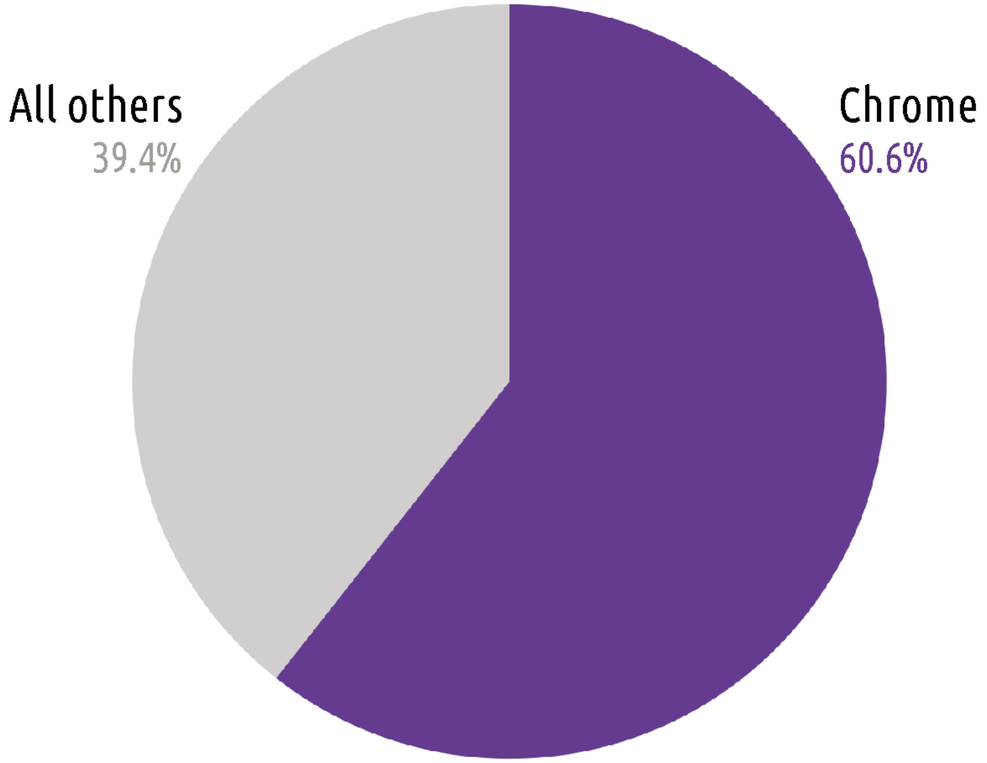
data processing

only certain types of data can fit into certain chart types, and vice versa

D

Chrome Dominates a Cluttered Browser Market

At September 2018, data source: gs.statcounter.com/browser-market-share



Stage 2

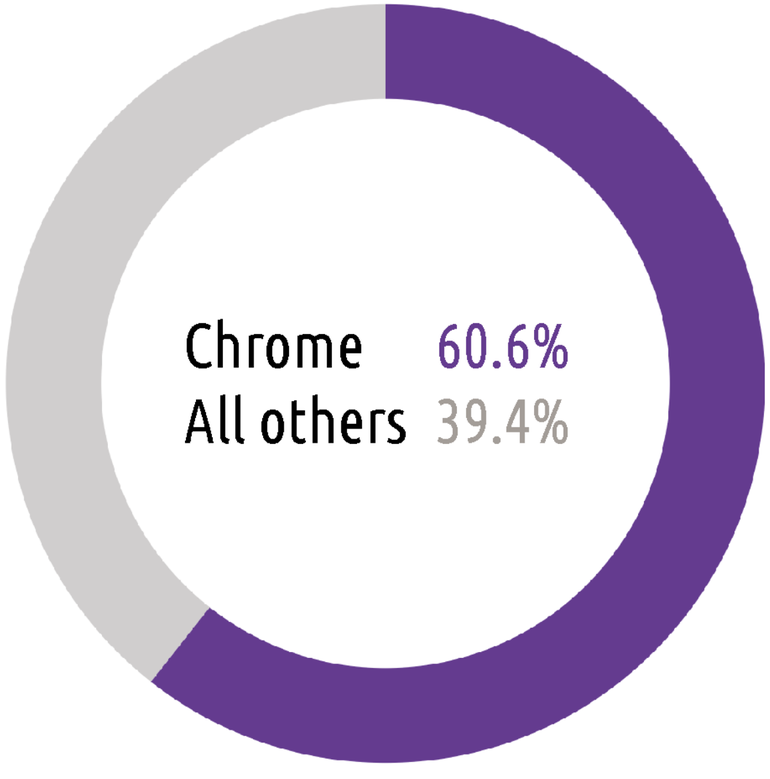
Working with **data**

data processing

only certain types of data can fit into certain chart types, and vice versa

E Chrome Dominates a Cluttered Browser Market

At September 2018, data source: gs.statcounter.com/browser-market-share



Stage 3

Establishing your **editorial** thinking

angle

choosing the angles of analysis dictates which chart type might be most relevant following the *chrts* taxonomy

treat every representation challenge on its own merits: having spatial data does not mean you must use a map

if regional/spatial information are not essential, the map composition may hinder the analysis rather than helping it

Principle 1

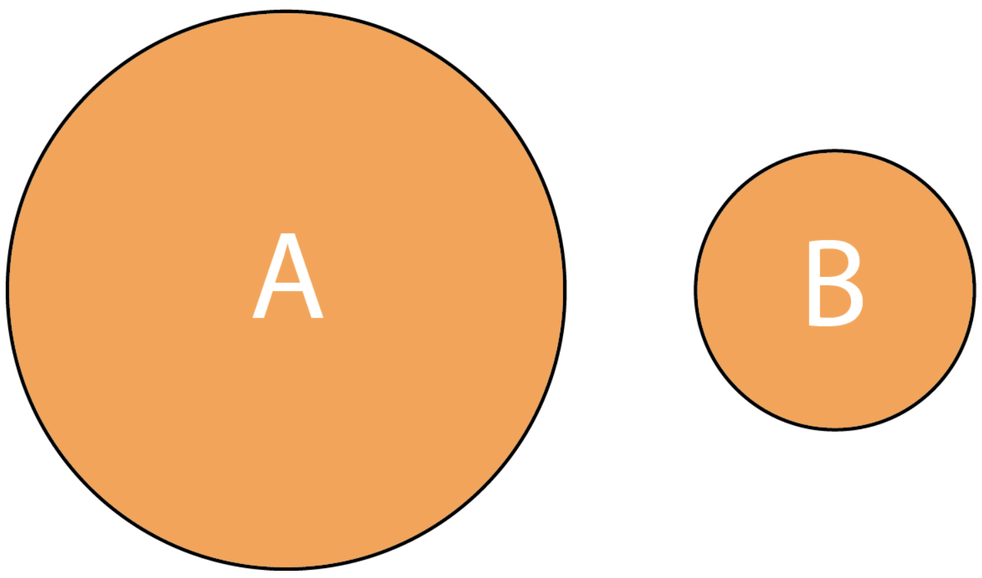
Good data visualisation is **TRUSTWORTHY**

trustworthyde sign

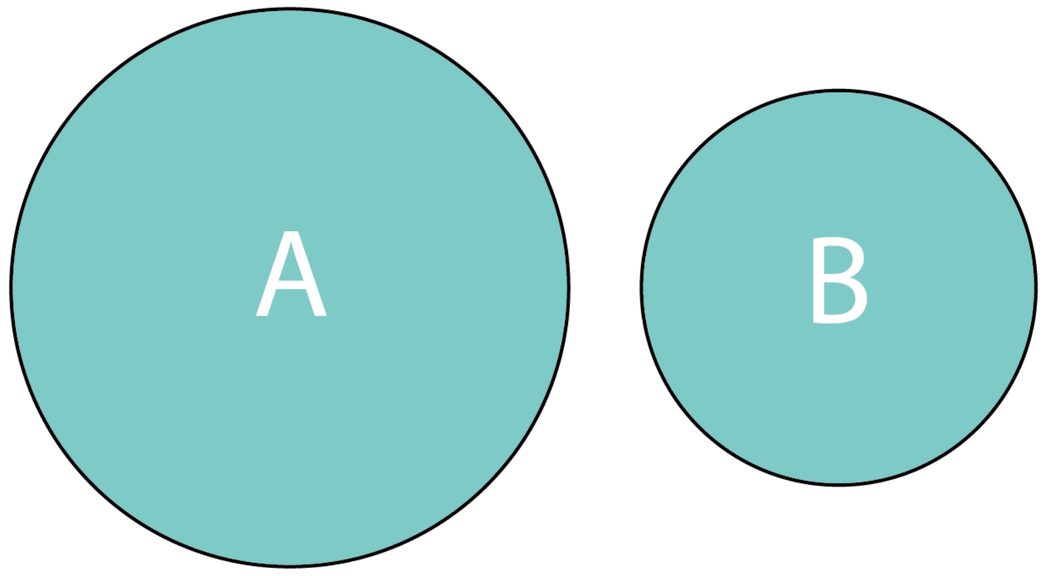
avoiding deception

geometric distortions

Variation in **diameter**



Variation in **area**



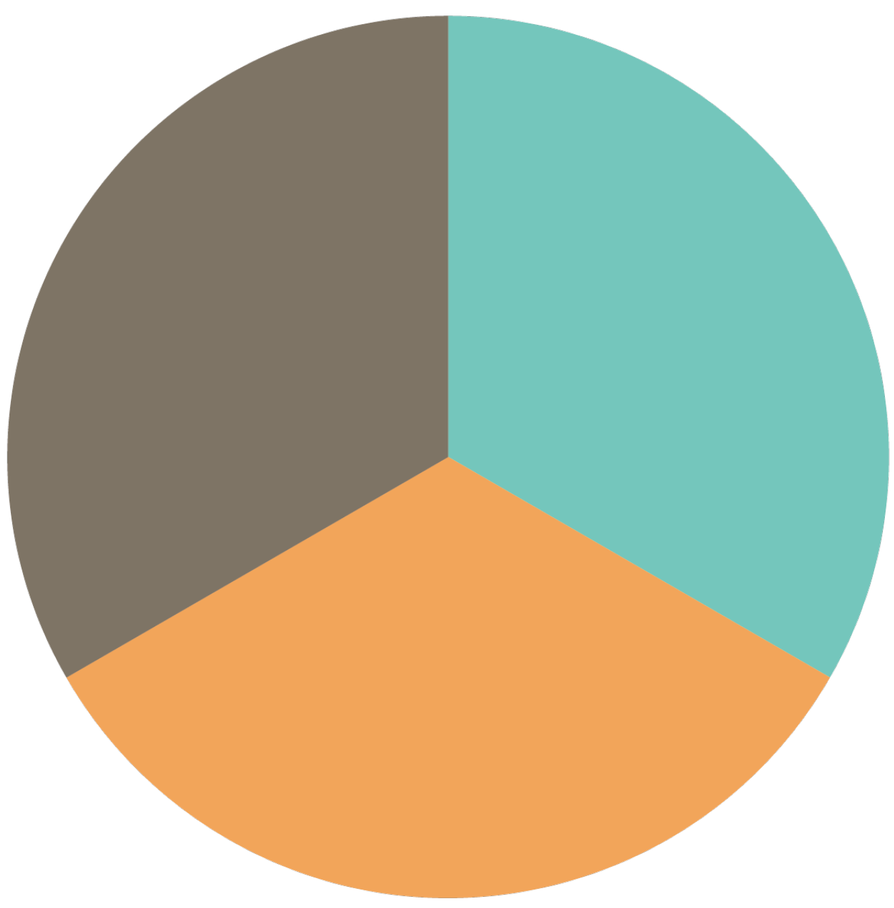
Principle 1

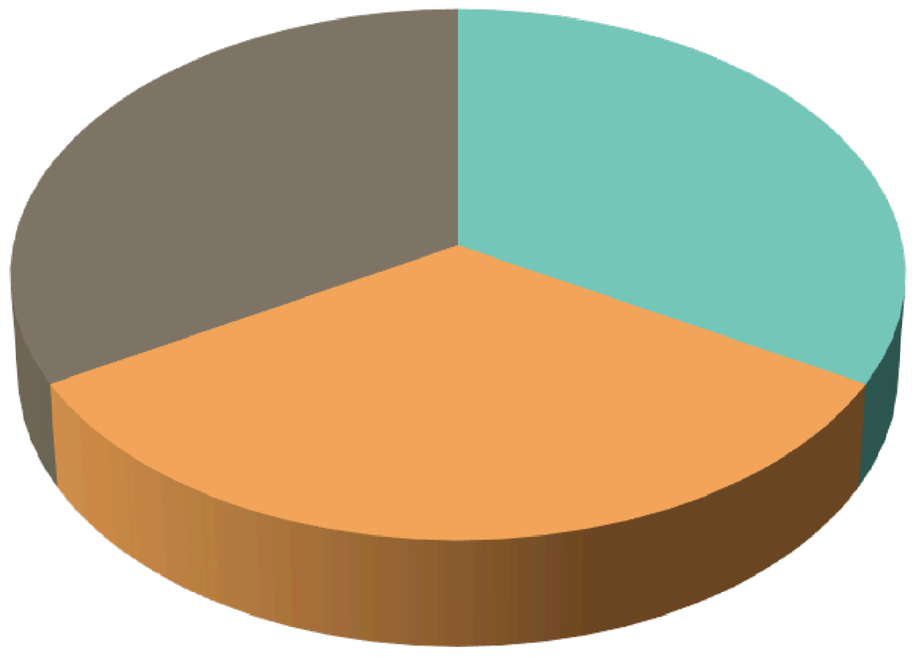
Good data visualisation is **TRUSTWORTHY**

trustworthyde sign

avoiding deception

3d decorative distortions





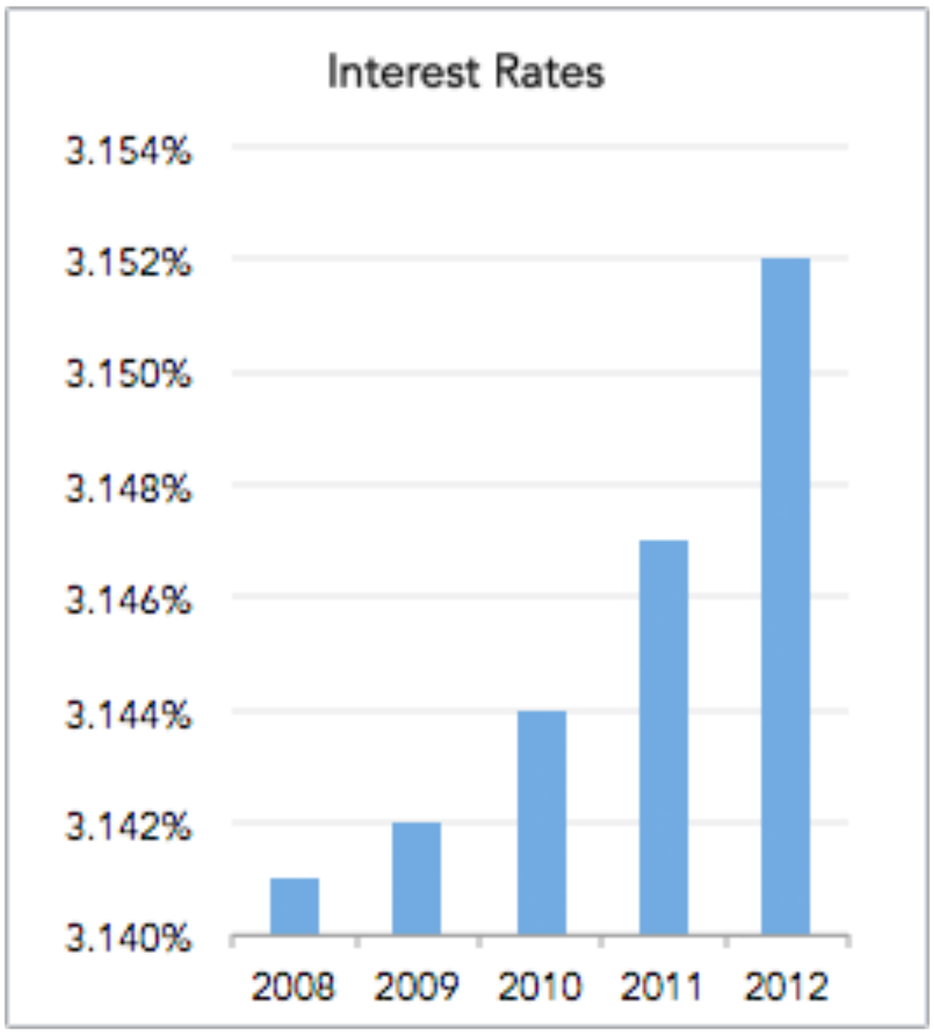
Principle 1

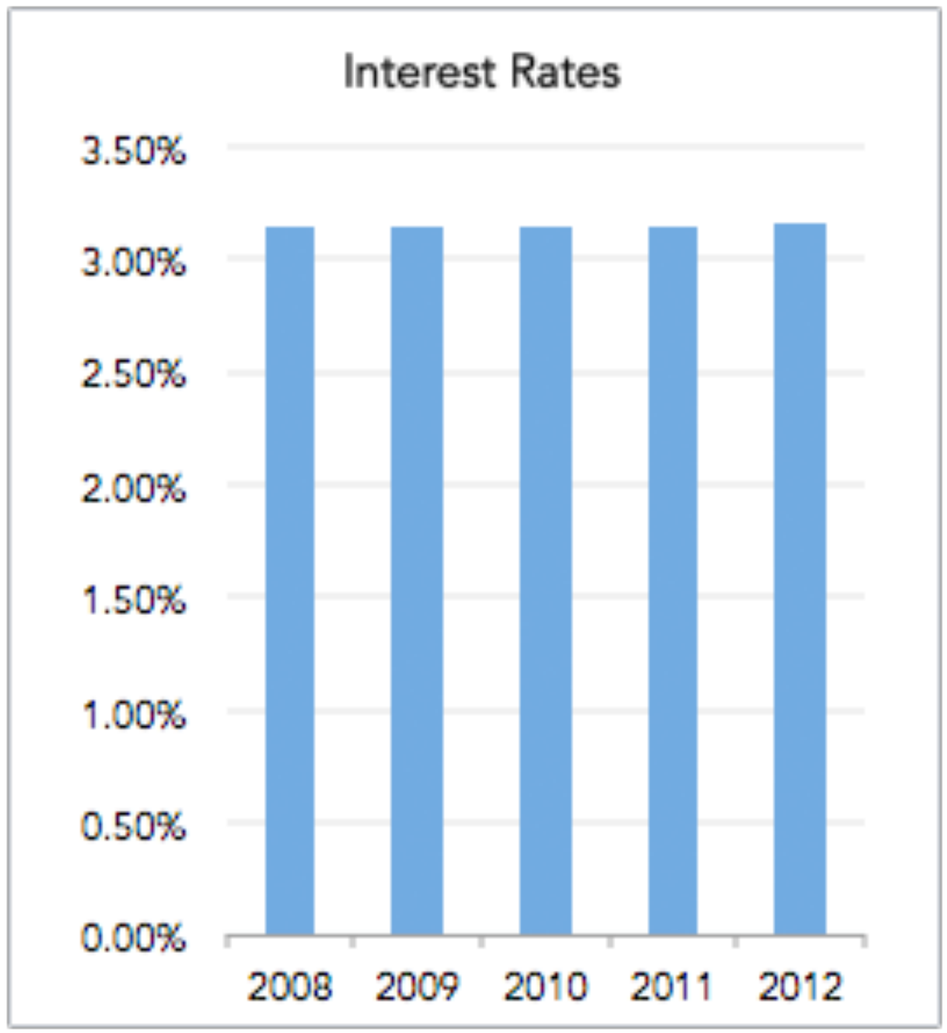
Good data visualisation is **TRUSTWORTHY**

trustworthyde sign

avoiding deception

truncated axes





Principle 2

Good data visualisation is **ACCESSIBLE**

accessible design

encoded overlays

incorporation of additional details:

- to explain further the context of values
- to amplify the interpretation of good/bad, normal/exceptional

they are *not* just annotations: they represent data values and require encoding choices

Principle 2

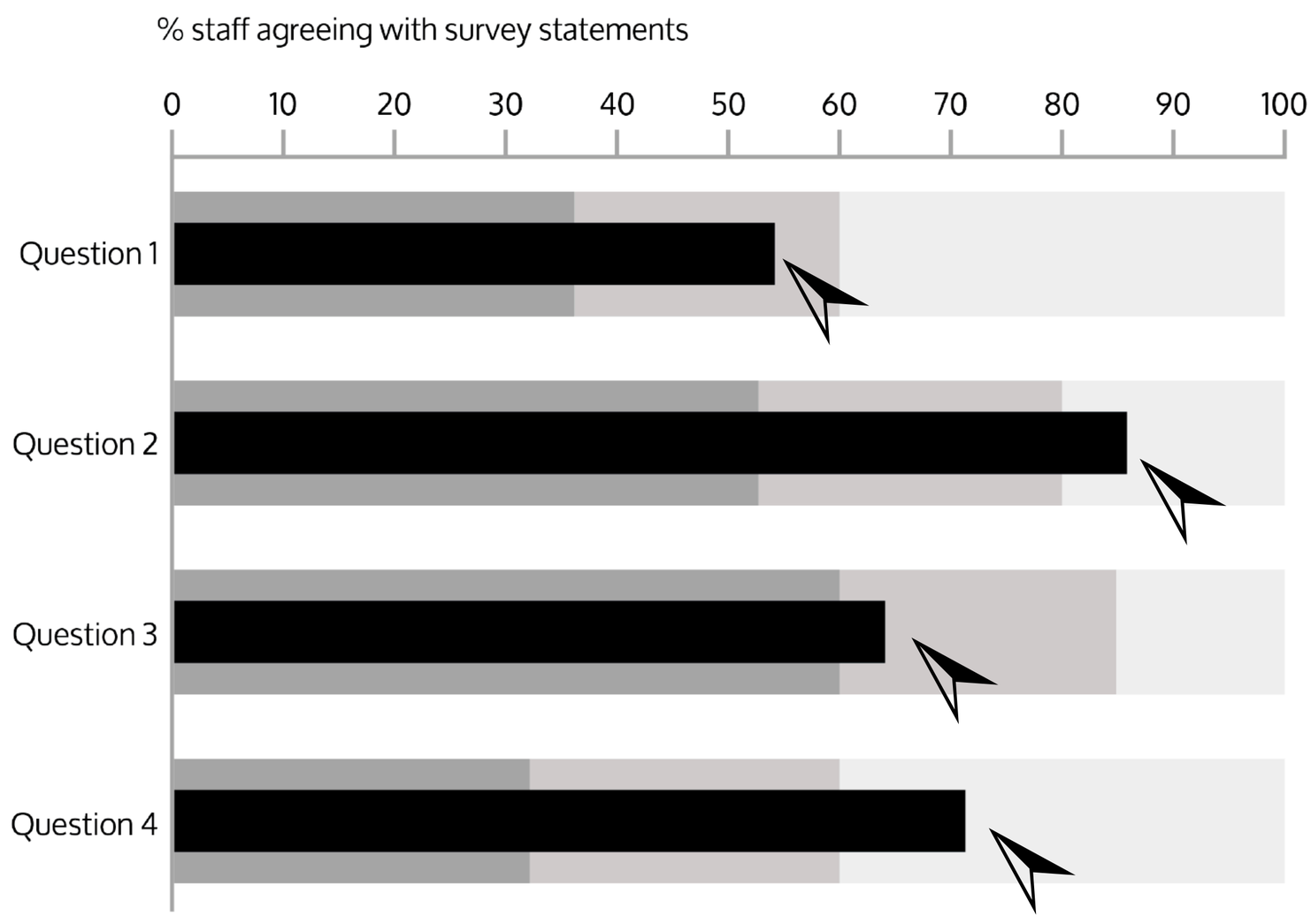
Good data visualisation is **ACCESSIBLE**

accessible design

encoded overlays

bandings

indicate the contrast between the main data value marks and the contextual judgement of historic or expected values



Principle 2

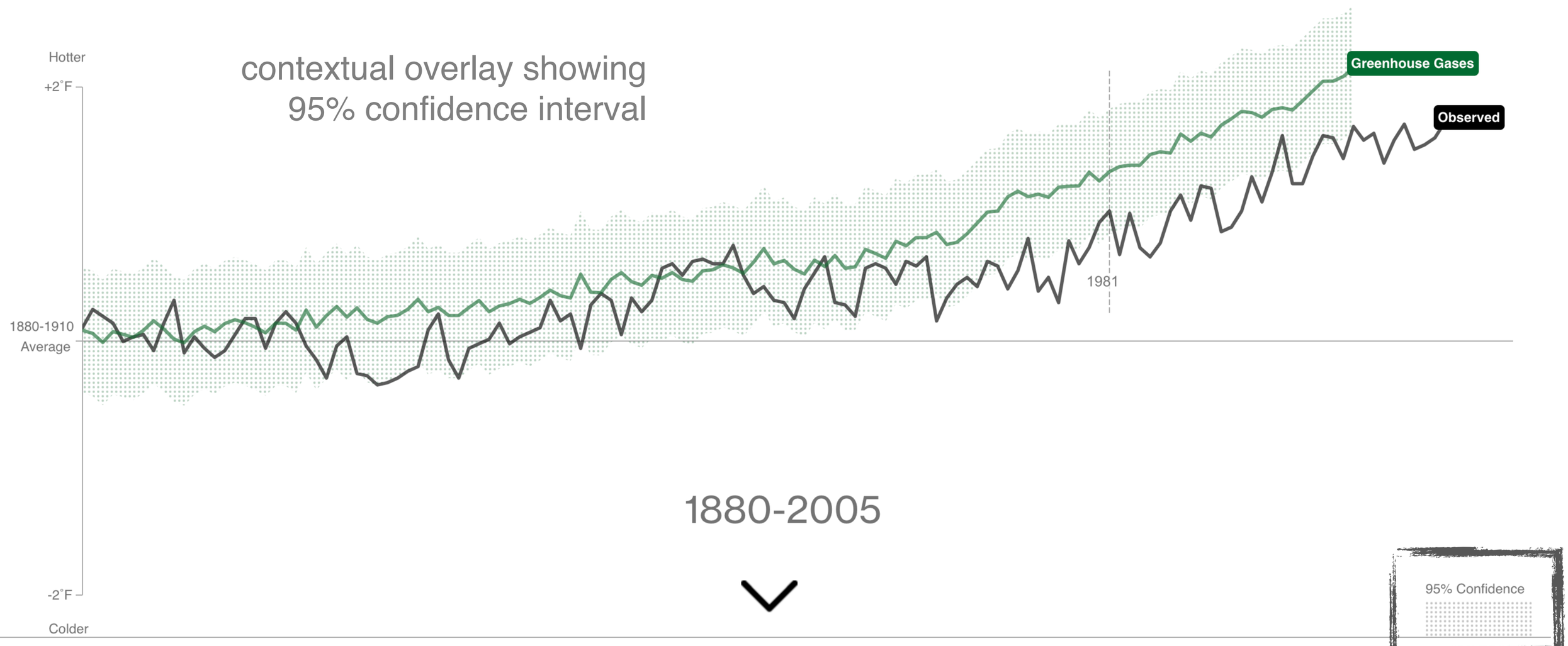
Good data visualisation
is **ACCESSIBLE**

accessible design

encoded overlays

bandings

contextual overlay showing
95% confidence interval



1880-2005

Principle 2

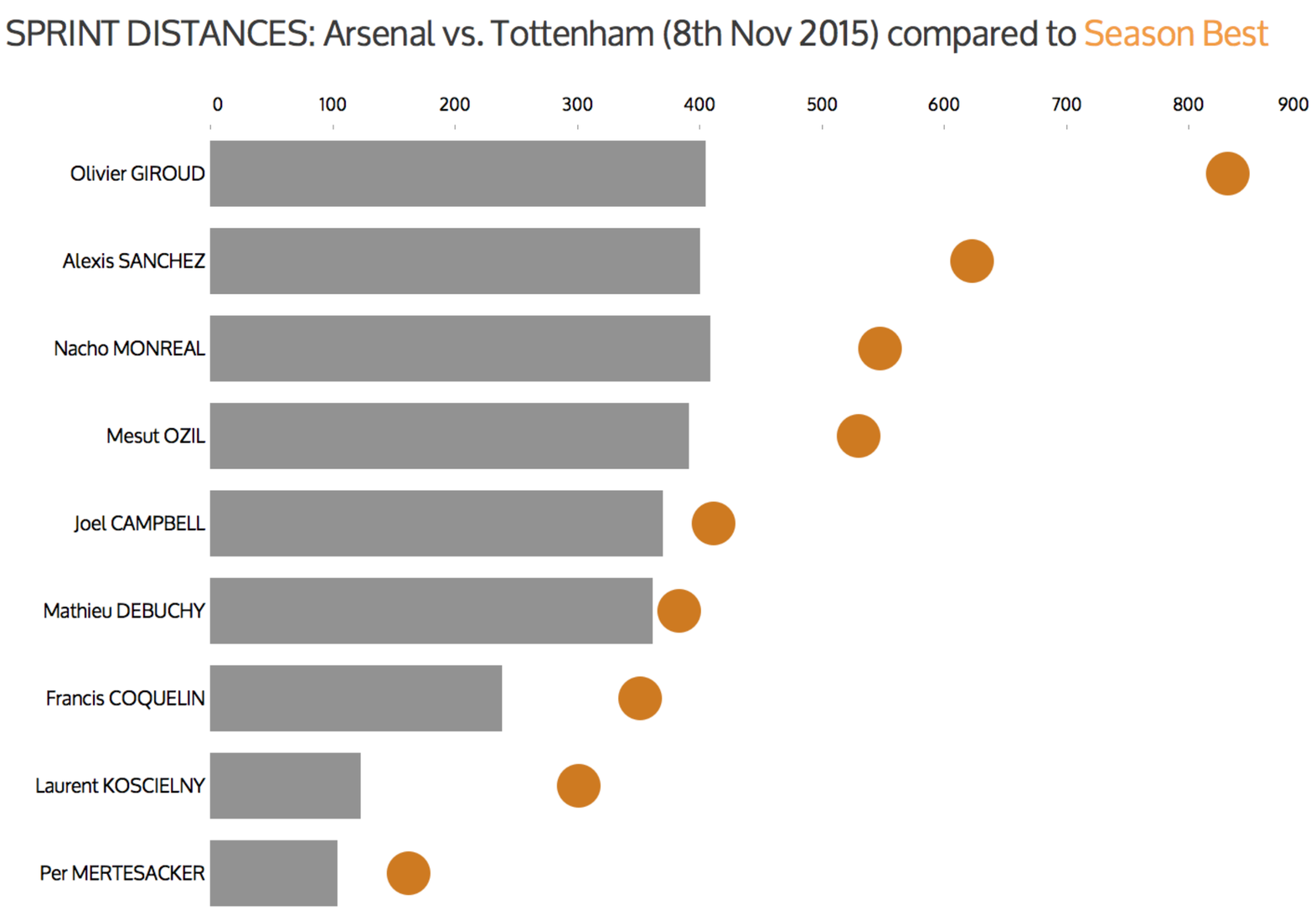
Good data visualisation is **ACCESSIBLE**

accessible design

encoded overlays

markers

additional points mark comparison versus a maximum value



Principle 2

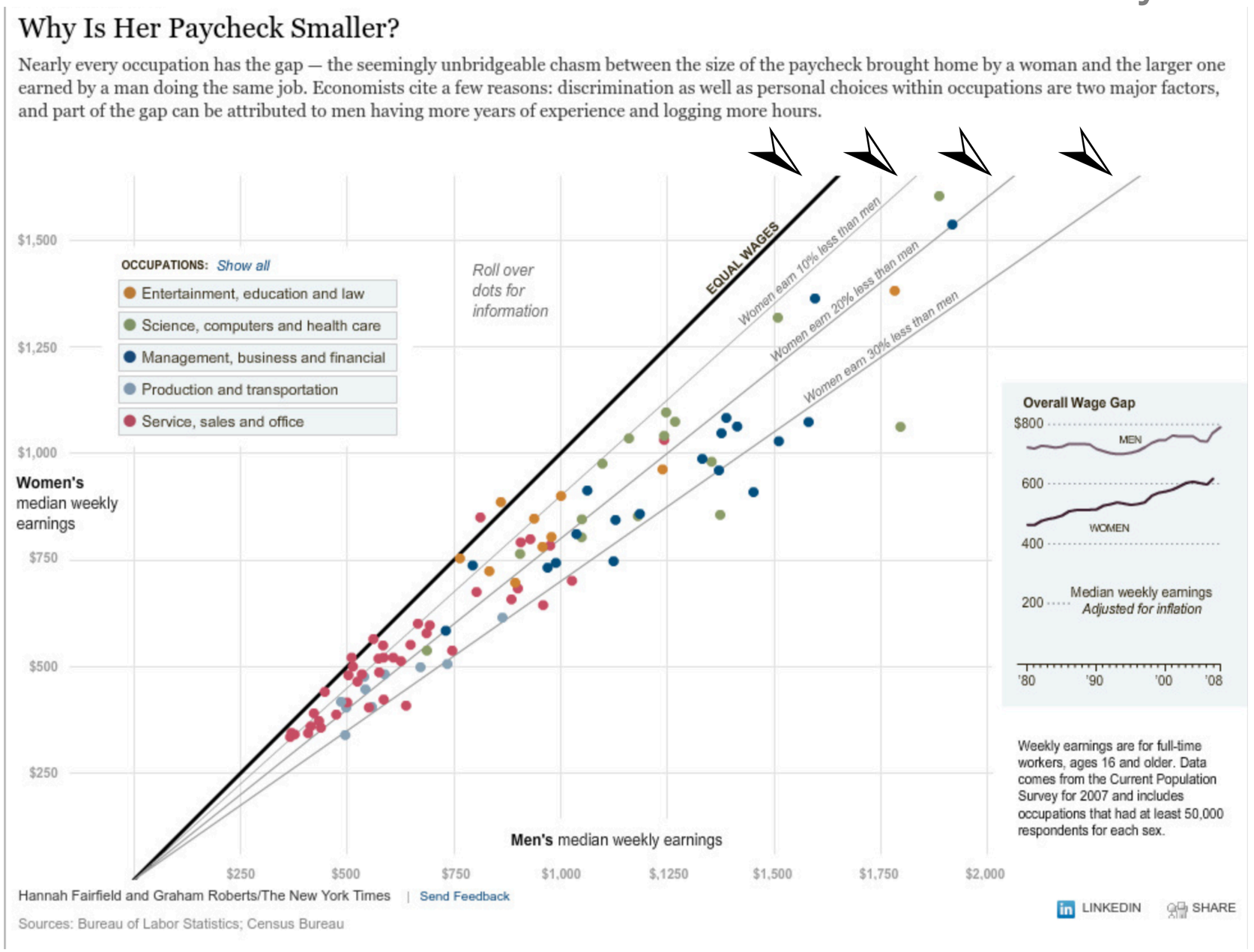
Good data visualisation is **ACCESSIBLE**

accessible design

encoded overlays

reference lines

in line charts and scatterplots, reference lines direct the eye towards calculated trends, constants, averages, correlations or best fits.



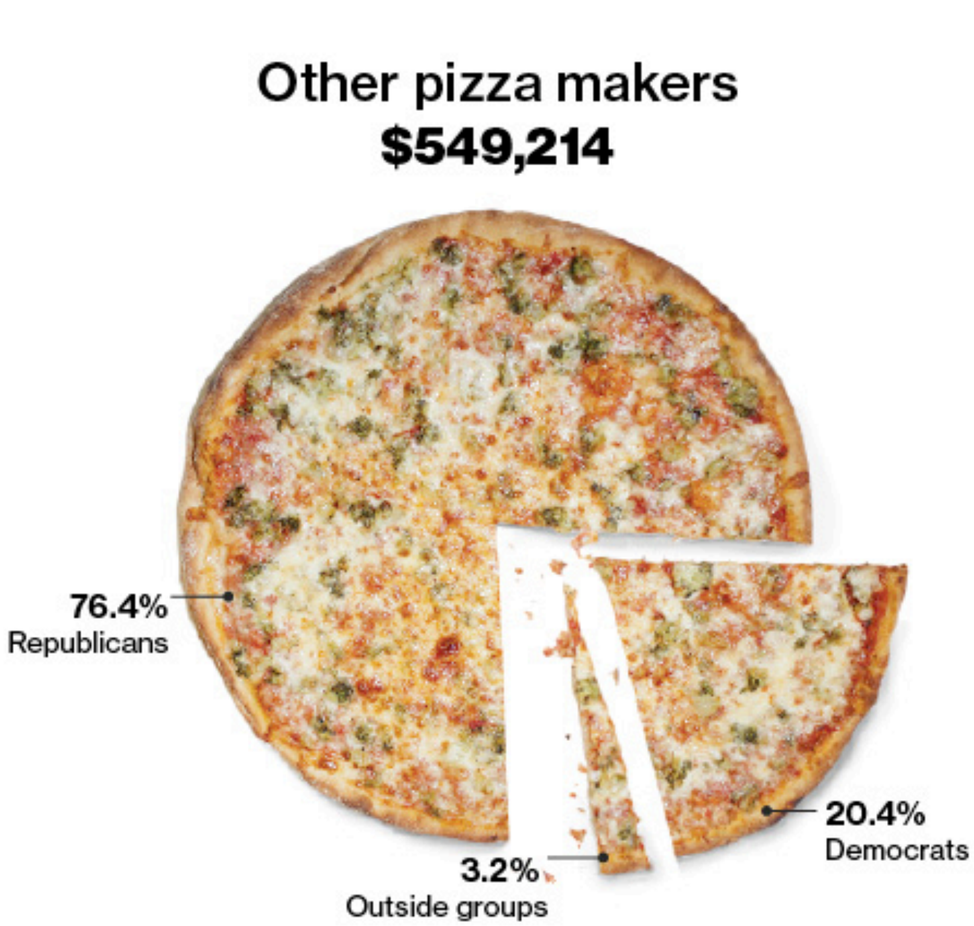
Principle 3

Good data visualisation is **ELEGANT**

elegant design

visual appeal

sometimes there might be scope in squeezing out an extra sense of stylistic association between visual and context



Stage 4

Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

composition

final layer of the design anatomy

making careful decisions about the physical attributes of an relationships between every visual property to ensure the optimum readability and meaning of the overall project

project composition
defining the layout and the hierarchy of the project

chart composition
defining the shape, size & layout choices for all components within the chart

project composition

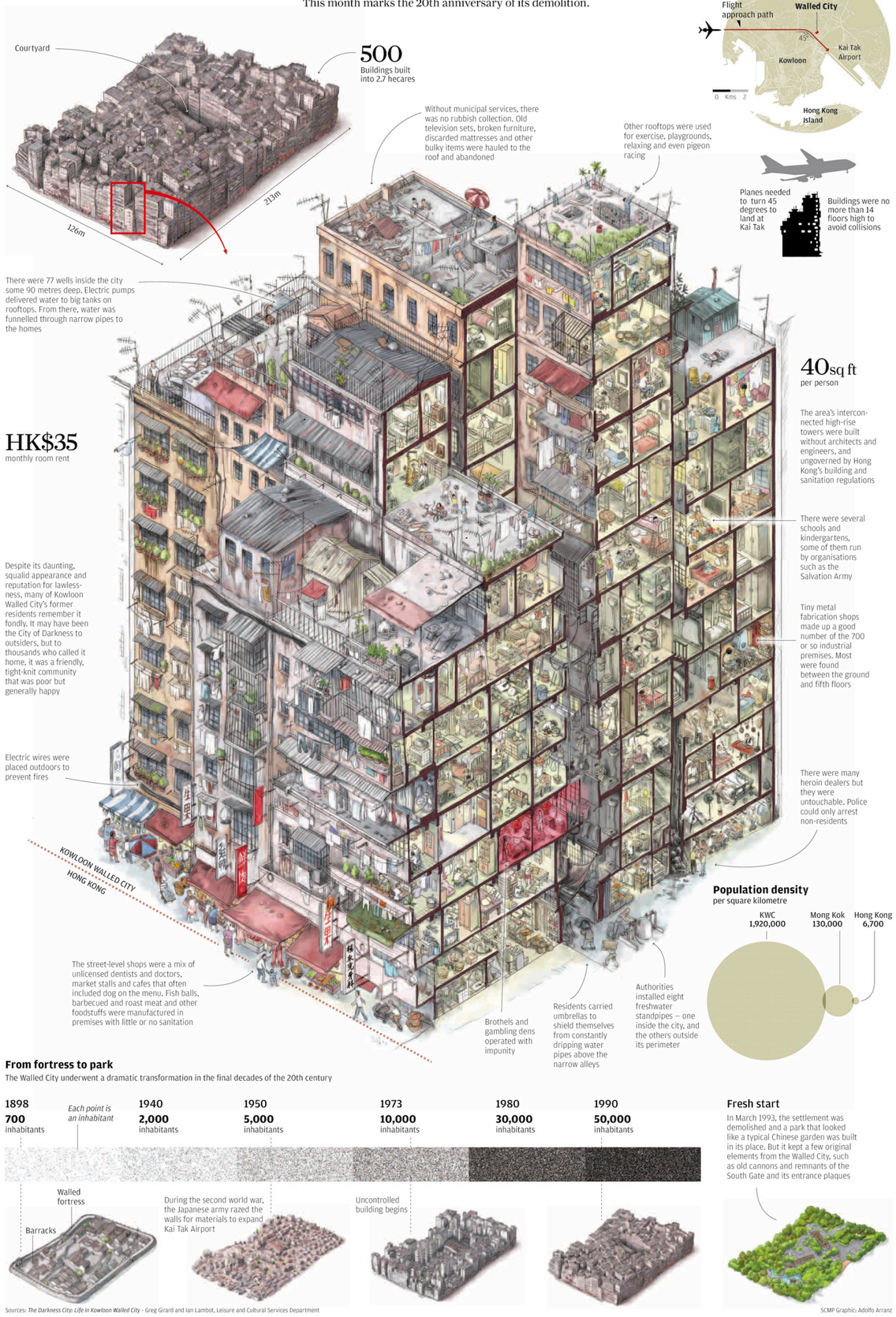
how to lay out and size all the visual components?

although established conventions do exist, this is usually an *iterative* process towards what *feels* like an optimal layout

hierarchy of content can be reached through careful choices in relative position & relative variation in size (and variation in colour for significance)

City of anarchy

Kowloon Walled City, located not far from the former Kai Tak Airport, was a remarkable high-rise squatter camp that by the 1980s had 50,000 residents. A historical accident of colonial Hong Kong, it existed in a lawless vacuum until it became an embarrassment for Britain. This month marks the 20th anniversary of its demolition.



Sources: *The Darkness City: Life in Kowloon Walled City* - Greg Girard and Ian Lambot, Leisure and Cultural Services Department

SCMP Graphic: Adolfo Arranz

Printed and published by South China Morning Post Publishers Ltd, Morning Post Centre, 22 Dai Fat Street, Tai Po Industrial Estate, Tai Po, Hong Kong. Tel: 2680 8888.

project composition

primary focal point

small thumbnail images for orientation

small supplementary illustrations at bottom for further information

kowloon walled city

Stage 4

Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

project composition

wireframing
sketching all layout/size across a single-page view,
including interactive functions

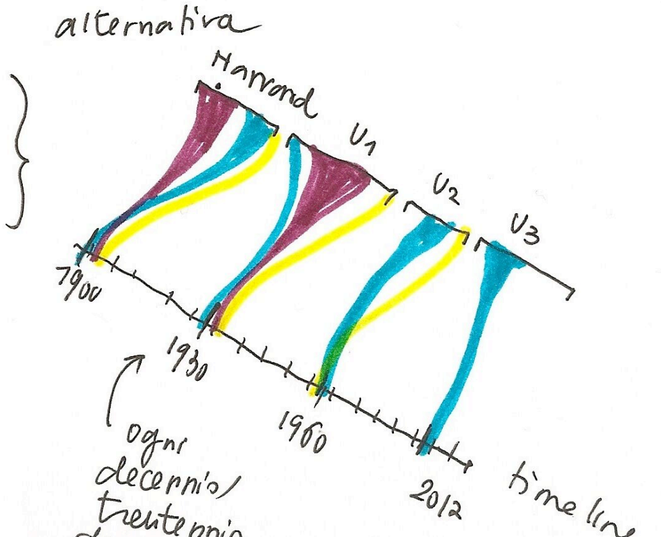
Stage 4

Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

project composition



wireframing example

Stage 4

Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

project composition

wireframing
sketching all layout/size across a single-page view,
including interactive functions

storyboarding
together with wireframing if the project is multipage

single pages included as cells in the big-picture hierarchy,
each one with its own wireframe

Stage 4

Developing your design **solution**

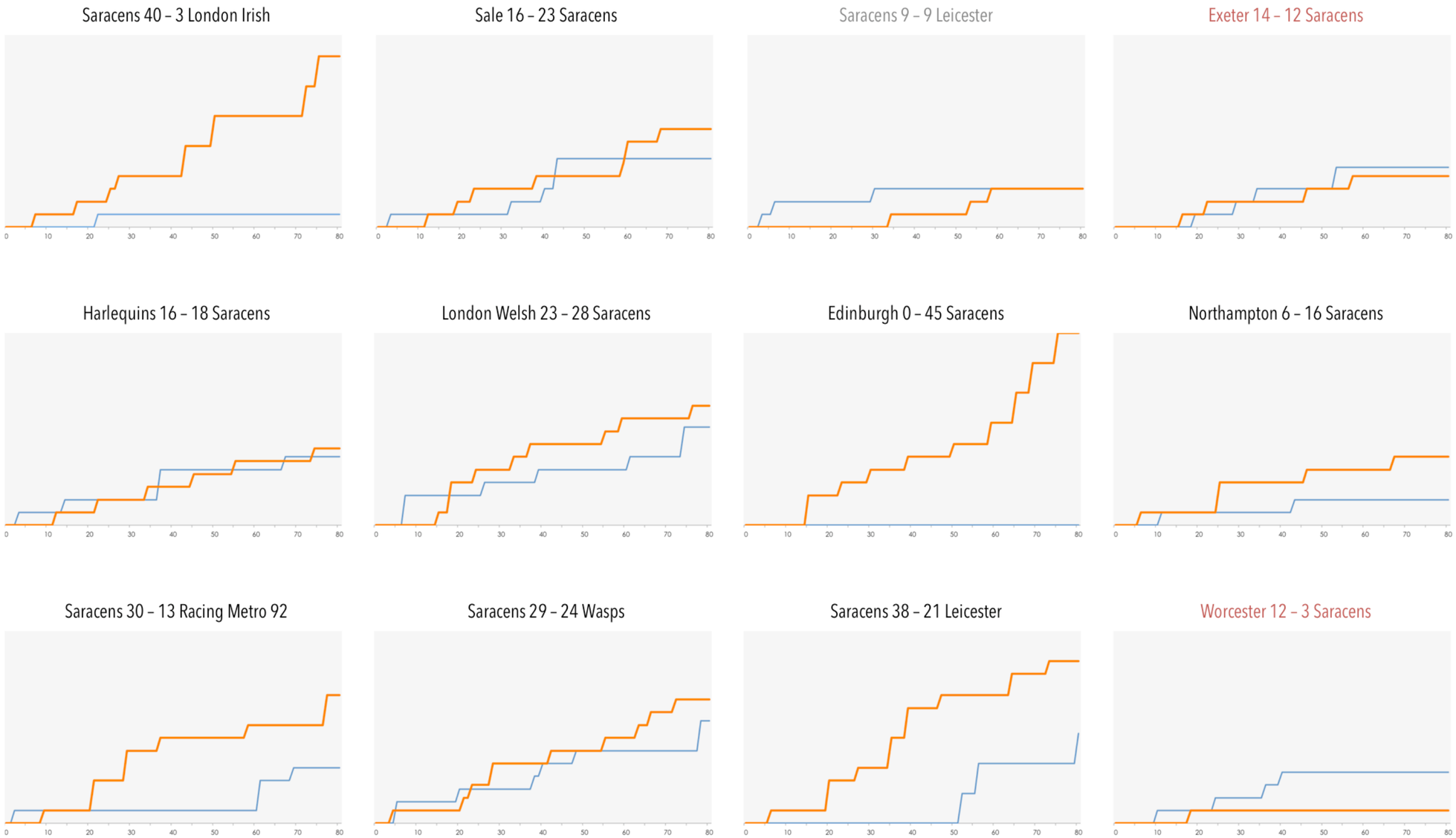
Stage 3

Establishing your **editorial** thinking

chart composition

chart size

do not be afraid of shrinking your chart, mind the font size



small multiples

Stage 4

Developing your design **solution**

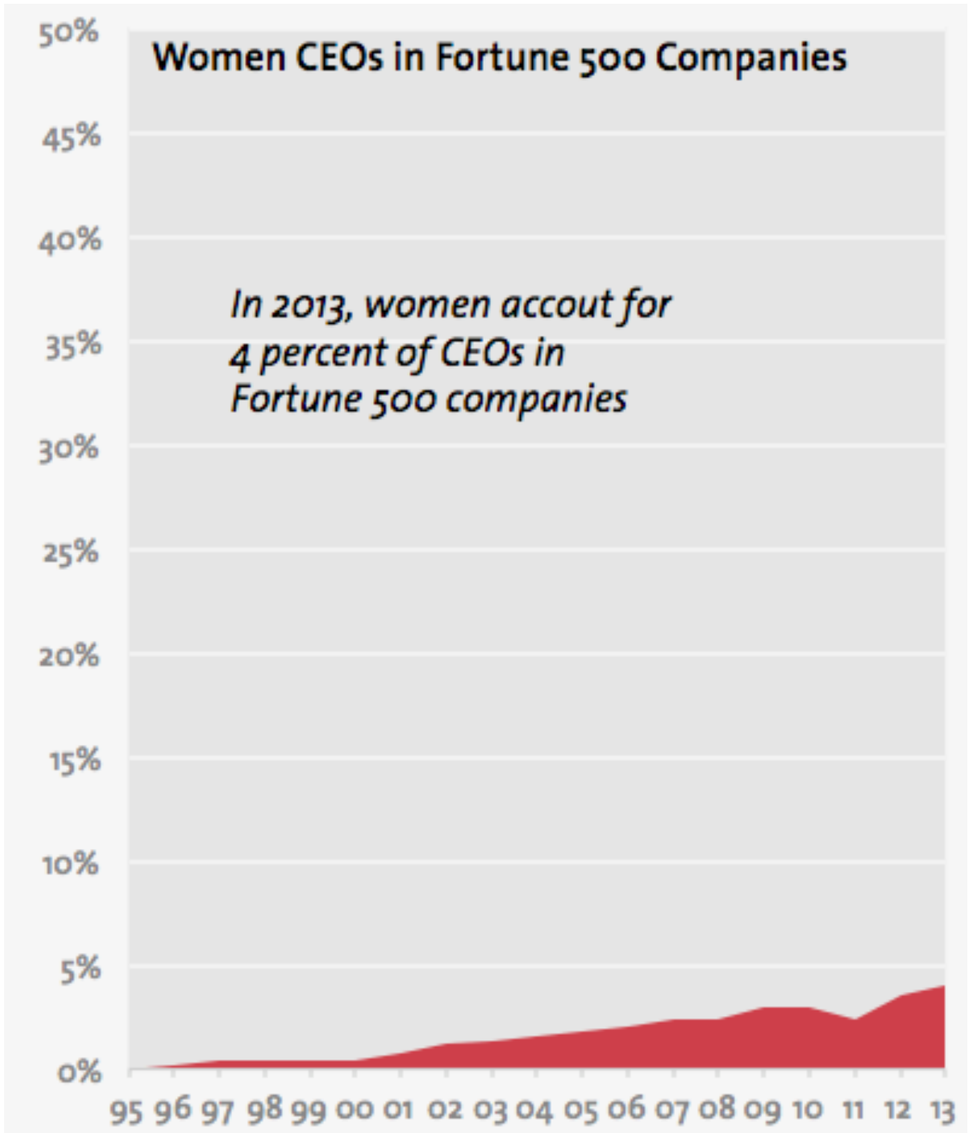
Stage 3

Establishing your **editorial** thinking

chart composition

chart scales

even the choice of scale can tell the viewer something



the apparently extreme scale marks the large disproportionality in women ceo’s amongst the fortune top-500 companies

Stage 4

Developing your design **solution**

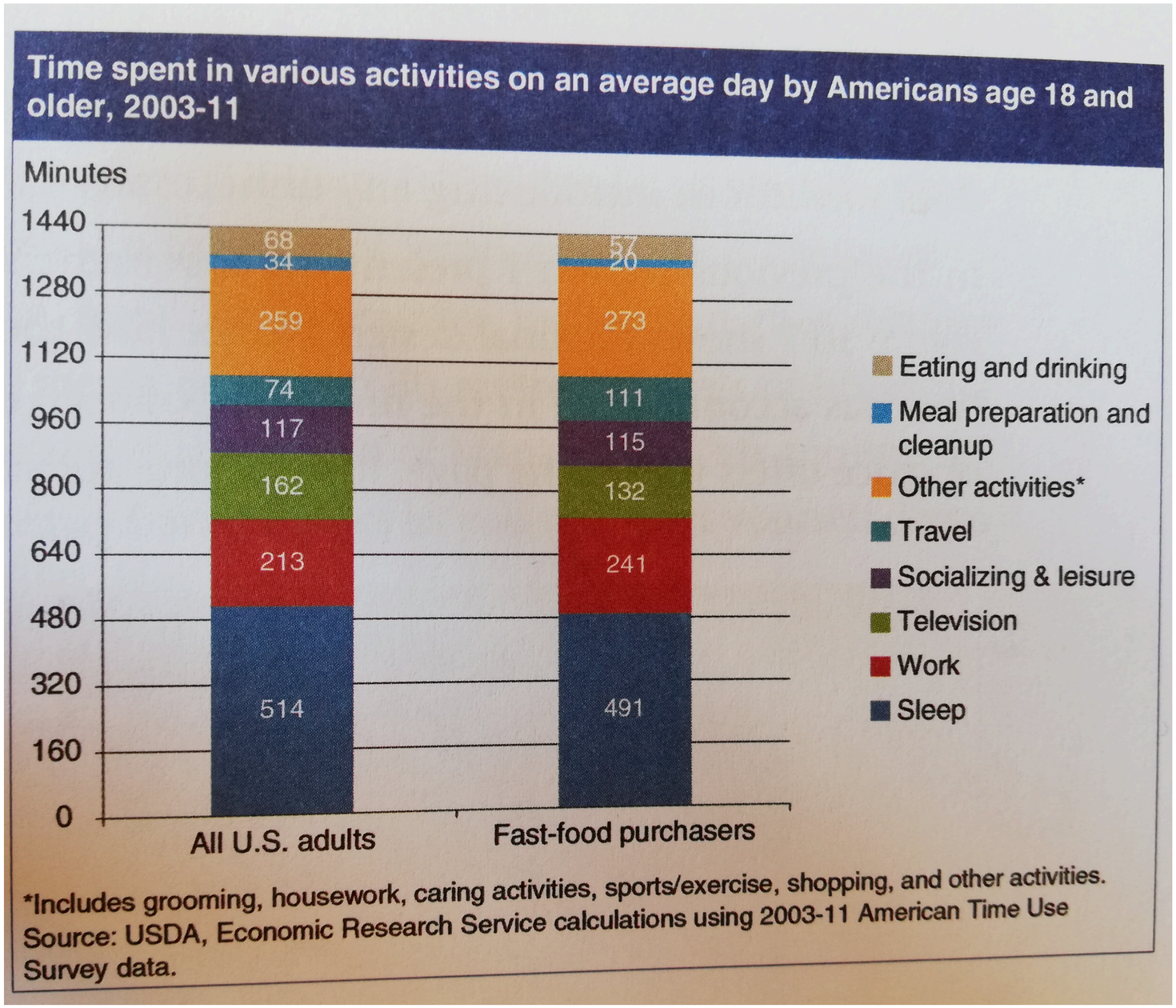
Stage 3

Establishing your **editorial** thinking

chart composition

chart scales

even the choice of scale can tell the viewer something



here the wrong scale choice has compromised readability — what's the point of dividing the 1440 minutes of a day in 160-mins intervals?

Stage 4

Developing your design **solution**

Stage 3

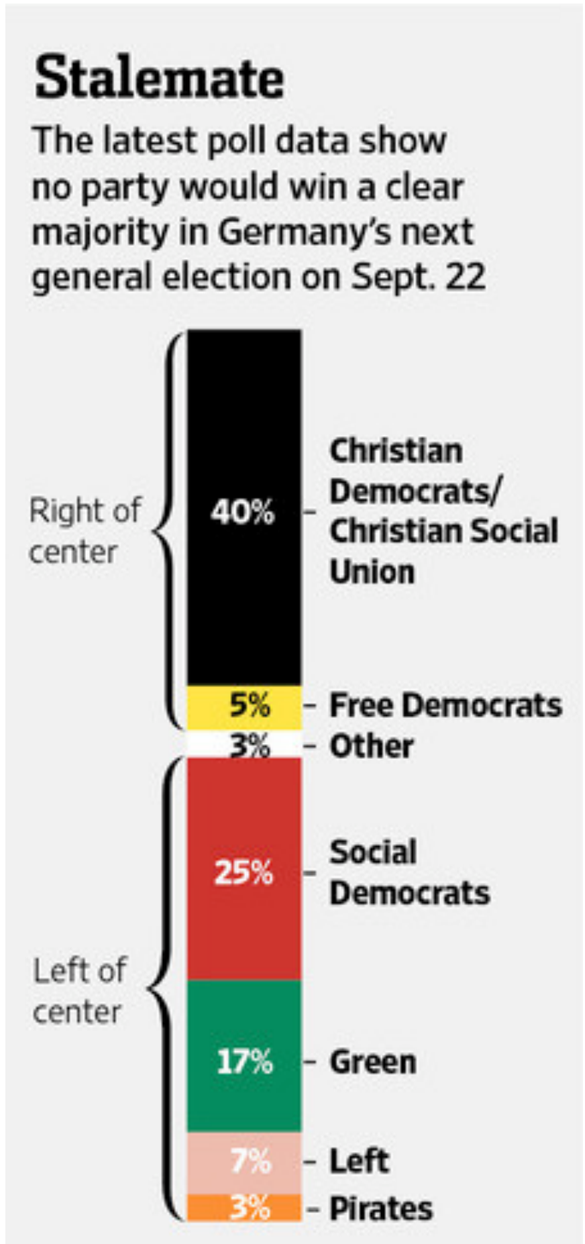
Establishing your **editorial** thinking

chart composition

chart orientation
adding an extra degree of readability

avoid label overlap

remember the Iraq's bloody toll



missing opportunity of using the left/ right duality

Stage 4

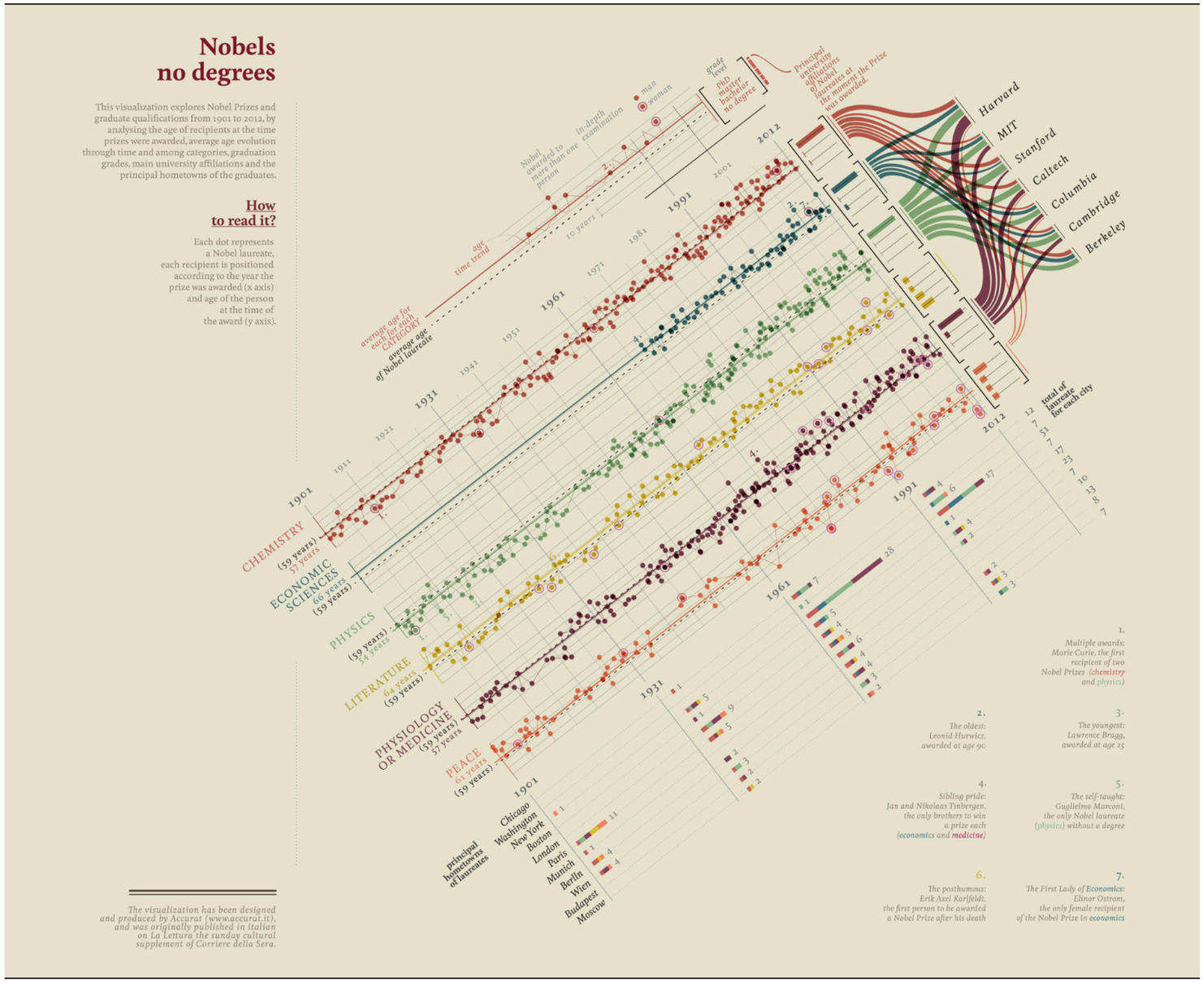
Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

chart composition

chart orientation
not only 90 degrees



final version of previous wireframe,
offering greater room in the page

Stage 4

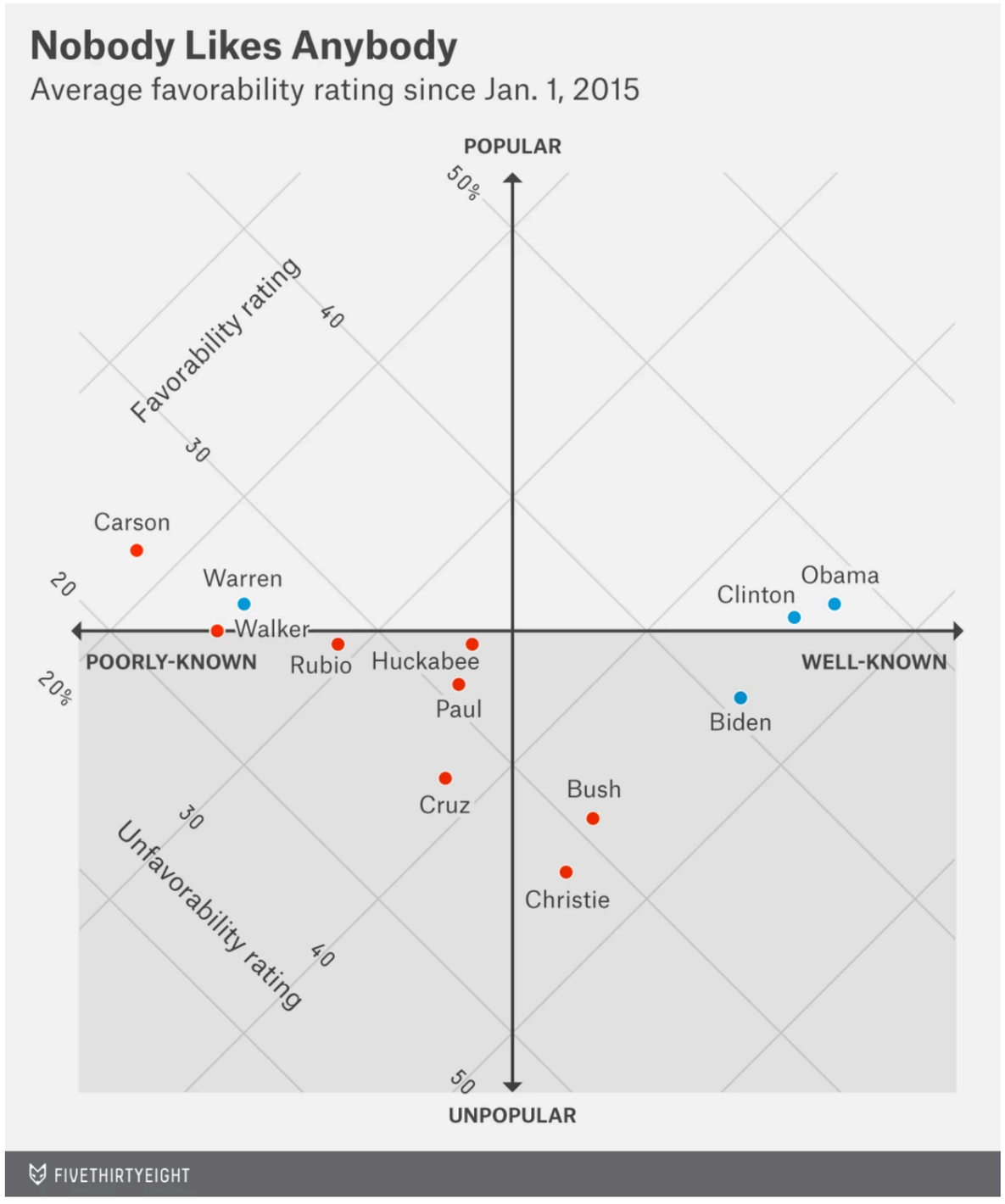
Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

chart composition

chart orientation
not only 90 degrees



45 degrees rotated scatterplot with 2x2 overplayed grid to make it easier to observe which values are located in each quadrant.

also emphasising the distinction between location at the top and bottom halves of the chart along the popularity axis, which is the primary focus of analysis

Stage 4

Developing your design **solution**

Stage 3

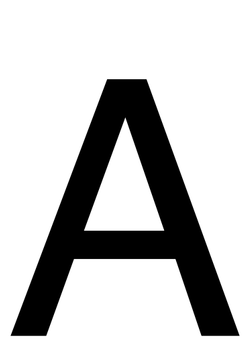
Establishing your **editorial** thinking

chart composition

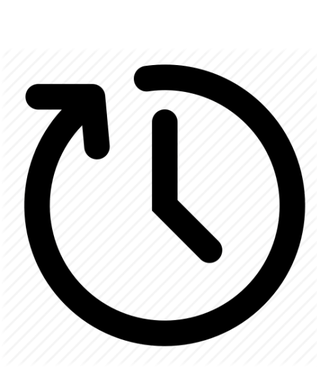
chart value sorting

latch rule

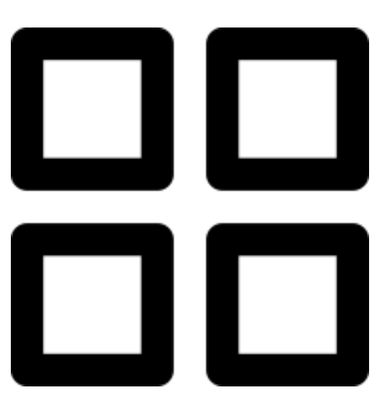
location



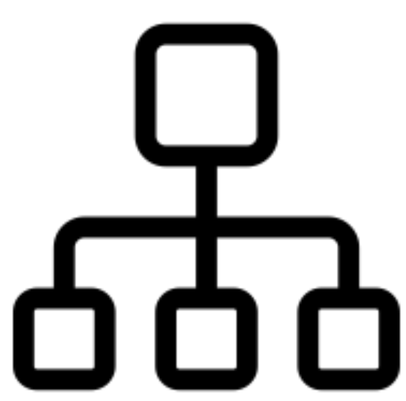
alphabet



time



category



hierarchy

Stage 4

Developing your design **solution**

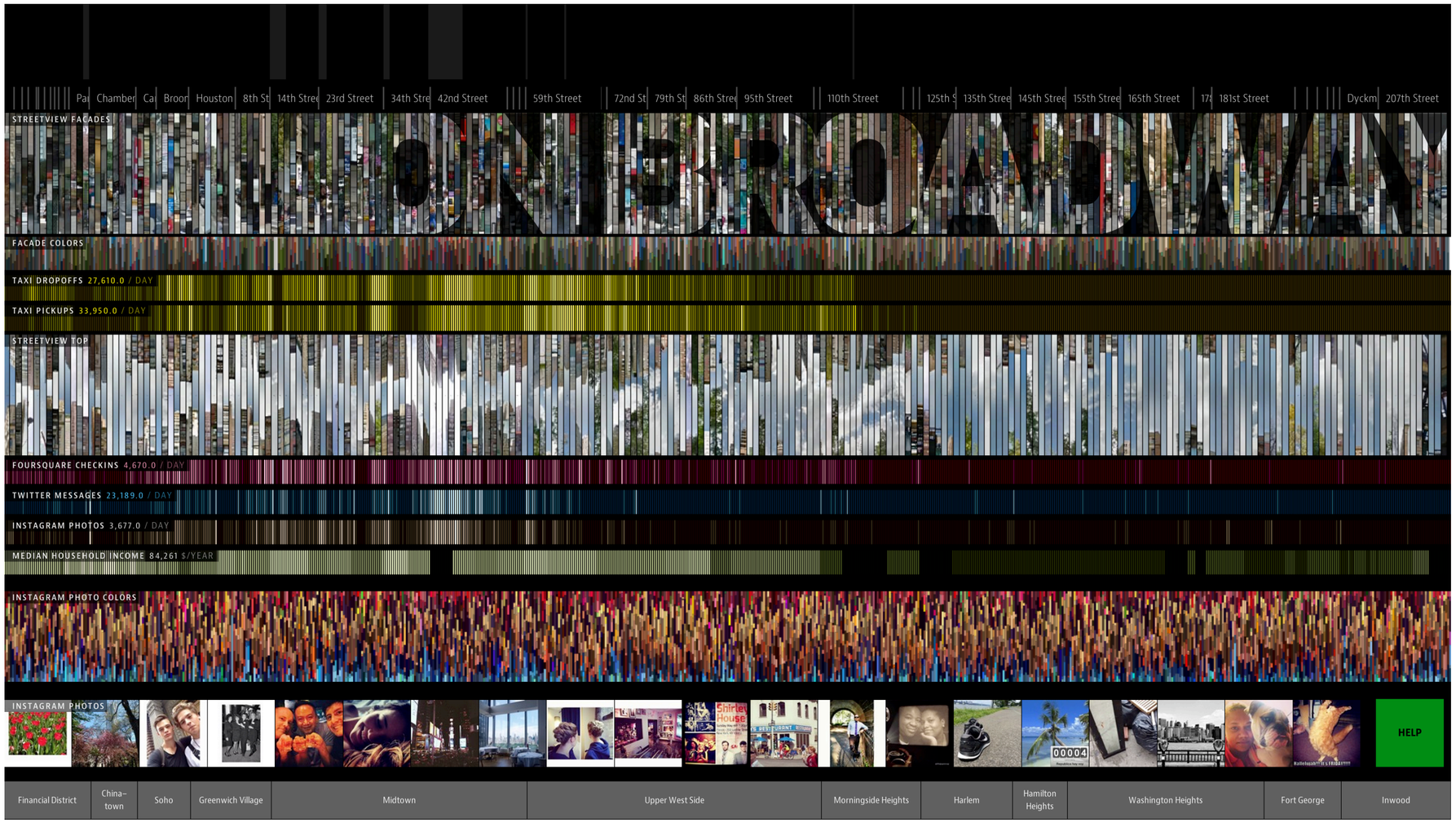
Stage 3

Establishing your **editorial** thinking

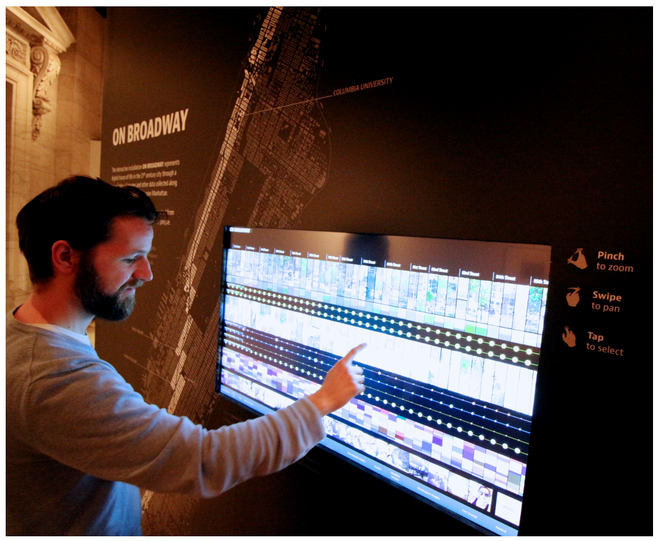
chart composition

chart value sorting

sequencing content according to spatial order
— only if offers the most logical sequence in content readability



on broadway installation



Stage 4

Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

chart composition

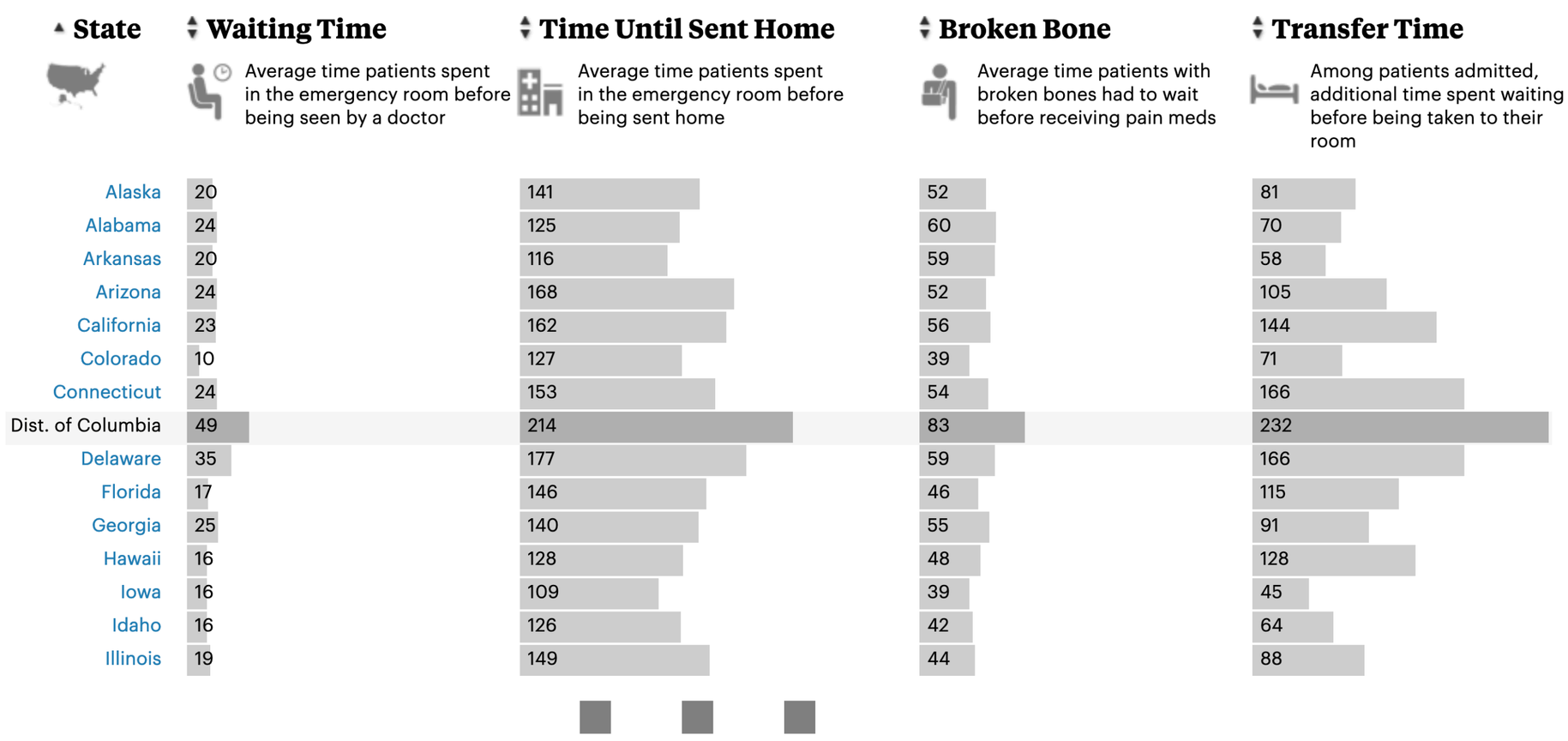
chart value sorting

A

alphabetical sorting facilitates efficient lookup & reference

requires user to rearrange items according to values

best sorting sequence if you do not want to imply any ranking



Stage 4

Developing your design **solution**

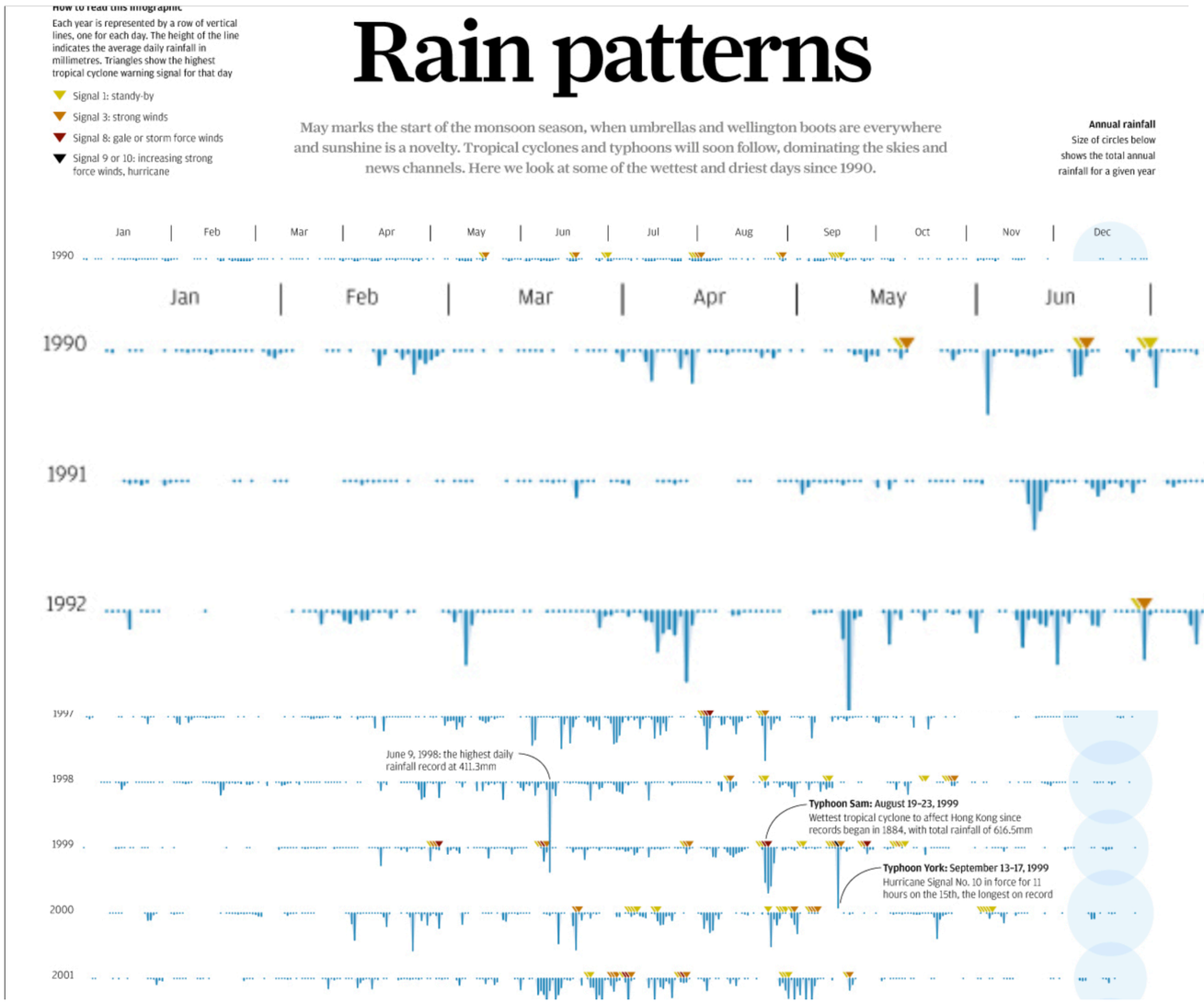
Stage 3

Establishing your **editorial** thinking

chart composition

chart value sorting

time-based sorting helps comparing changes over time



interest in seasonality of patterns makes chronological sorting meaningful

Stage 4

Developing your design **solution**

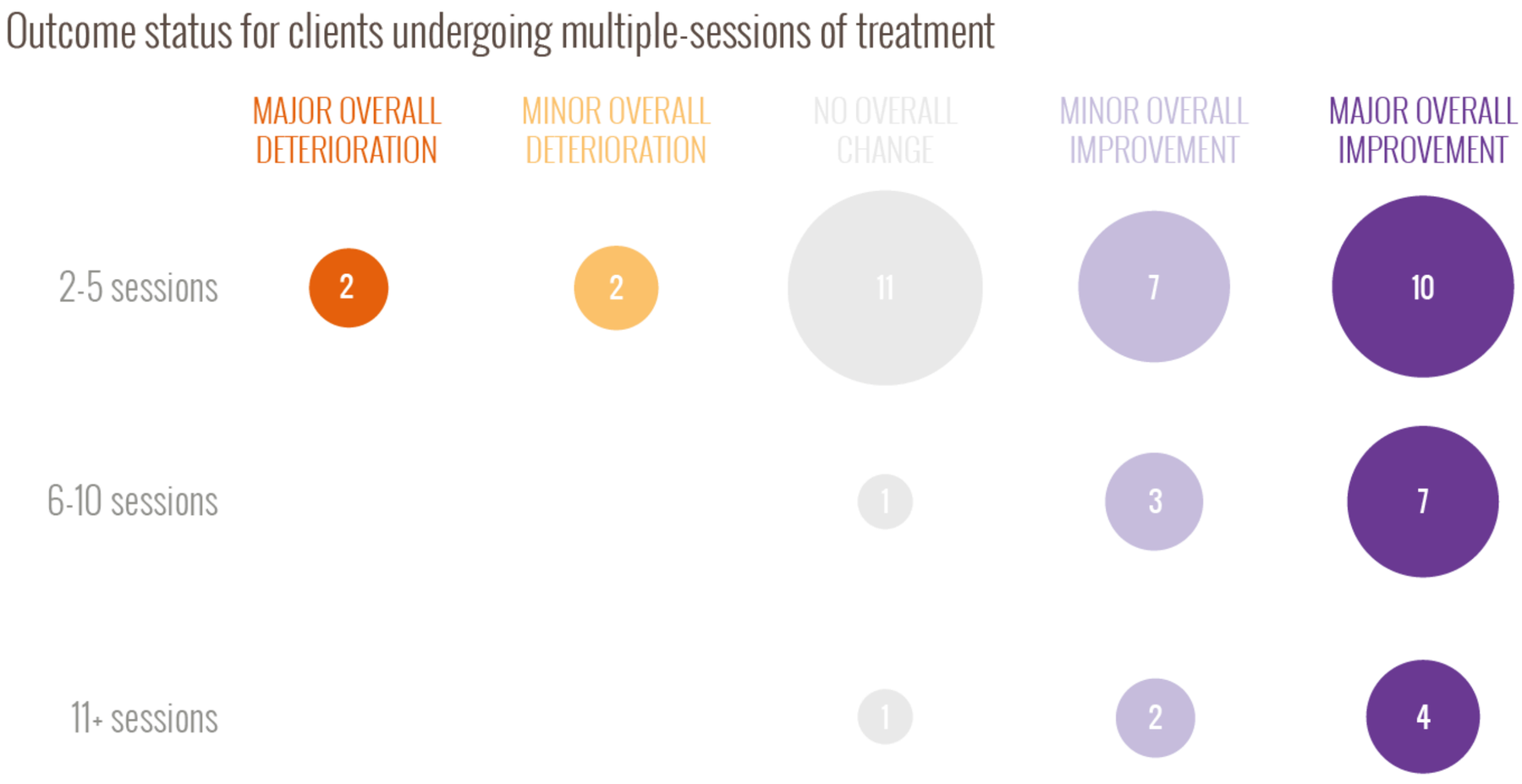
Stage 3

Establishing your **editorial** thinking

chart composition

chart value sorting

categorical sorting implies data is organised in categories ordered by a ranking inherited by their values or unique to the subject
(e.g., in ⚽, goalkeeper ➢ defenders ➢ midfielders ➢ attackers)



here colour is an example of redundant encoding; the color scheme was inherited by another panel of the dataviz

Stage 4

Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

chart composition

chart value sorting

hierarchical sorting is defined by increasing/decreasing quantities
efficient perception of size/distribution/ranking

Stage 4

Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

composition

influencing factors

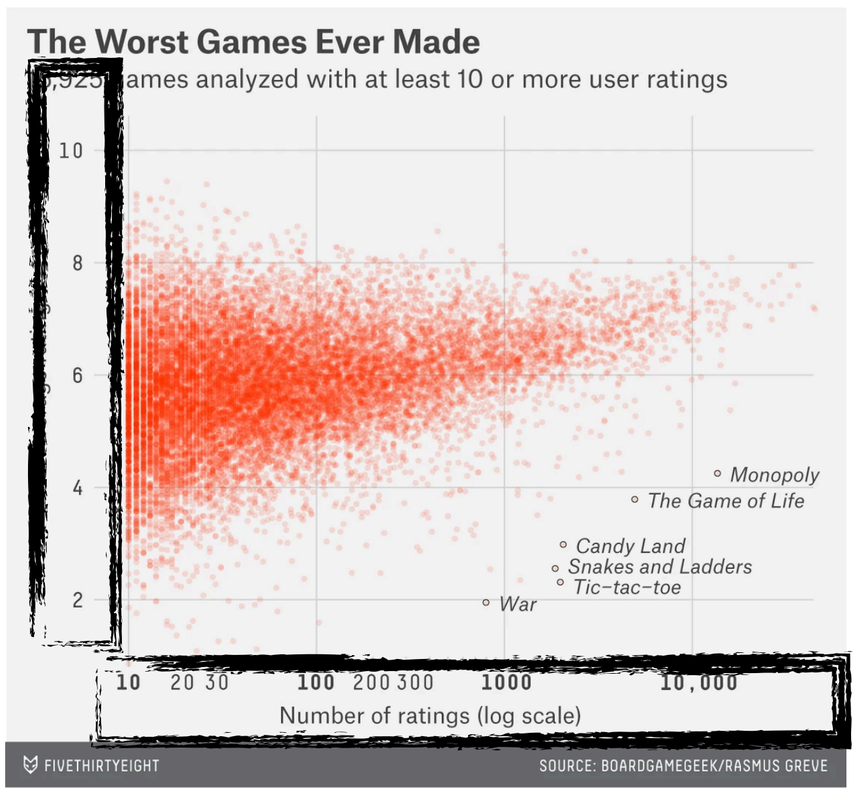
format

what is the shape/size of the primary format?
how transferable is the solution across different platforms?

data

how to legitimately fit data into the given canvas?

change scale to fit a square



change column style to fit data

Month on month inflation % 2000–2008

	2000	2001	2002	2003	2004	2005	2006	2007	2008
Jan	55	57	116	208	628	133	613	1593	100,580
Feb	48	57	116	220	602	127	782	1720	165,000
Mar	50	55	113	228	583	123	913	2200	355,000
Apr	53	56	114	269	505	129	1092	3714	736,604
May	58	55	122	300	448	144	1193	4539	1,800,000
Jun	59	64	114	364	394	164	1184	7251	
Jul	53	70	123	399	362	254	993	7634	220,000,000
Aug	53	76	135	426	314	265	1204	6592	231,000,000
Sept	62	86	139	455	251	359	1023	7892	
Oct	60	97	144	525	209	411	1070	14840	
Nov	56	103	175	619	149	502	1098	26470	
Dec	55	112	198	598	132	585	1281	66000	

Stage 4

Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

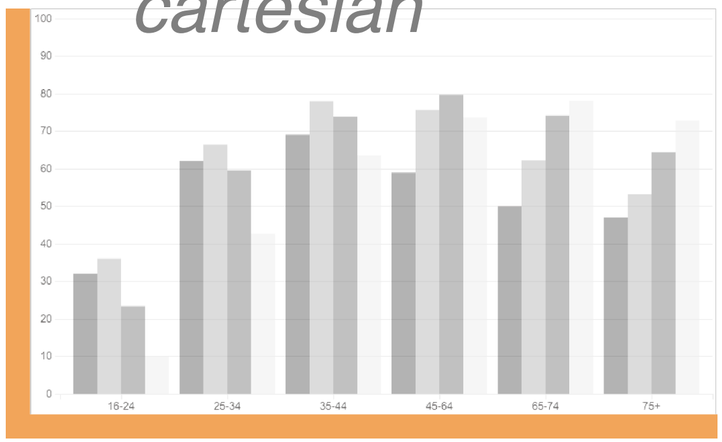
composition

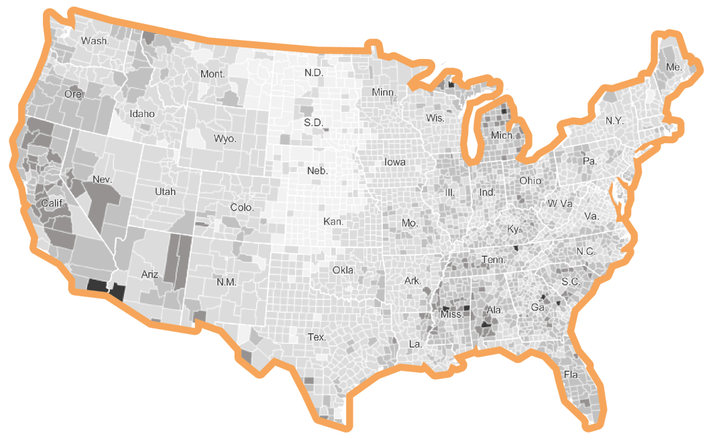
influencing factors

chart type

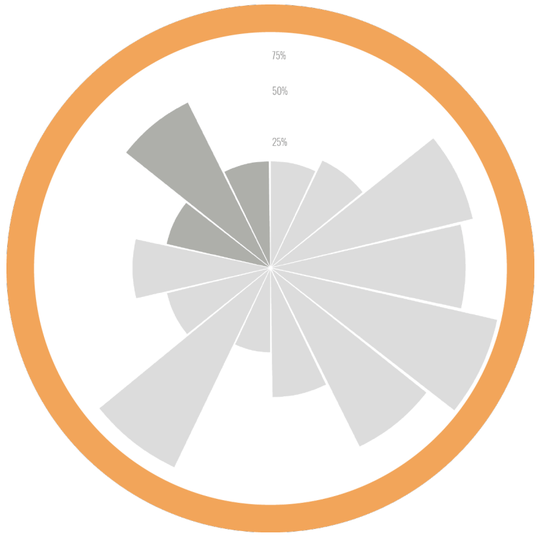
what are the spatial consequences of the chosen chart?

cartesian

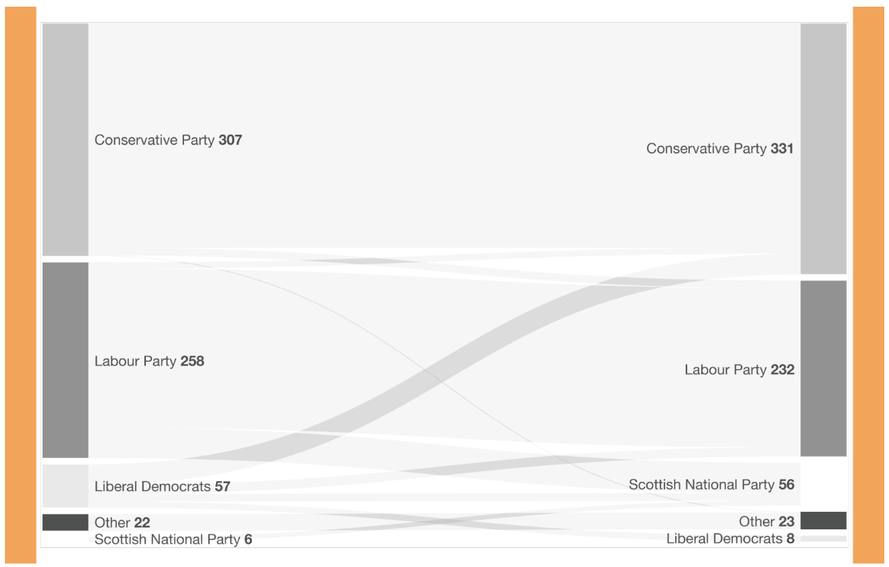




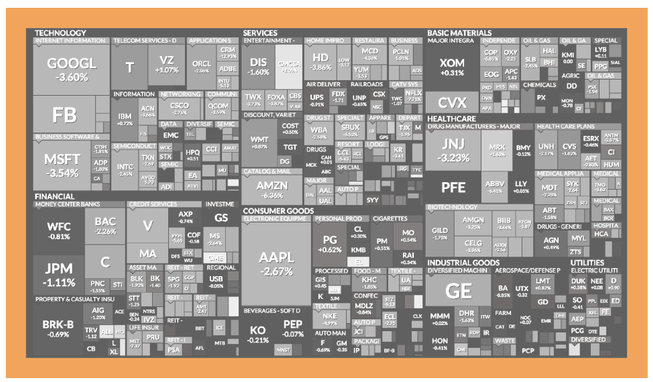
spatial



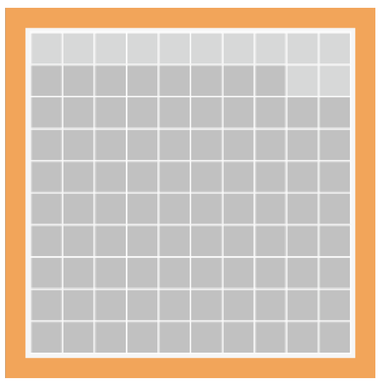
radial



columnar



enclosure



Stage 4

Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

composition

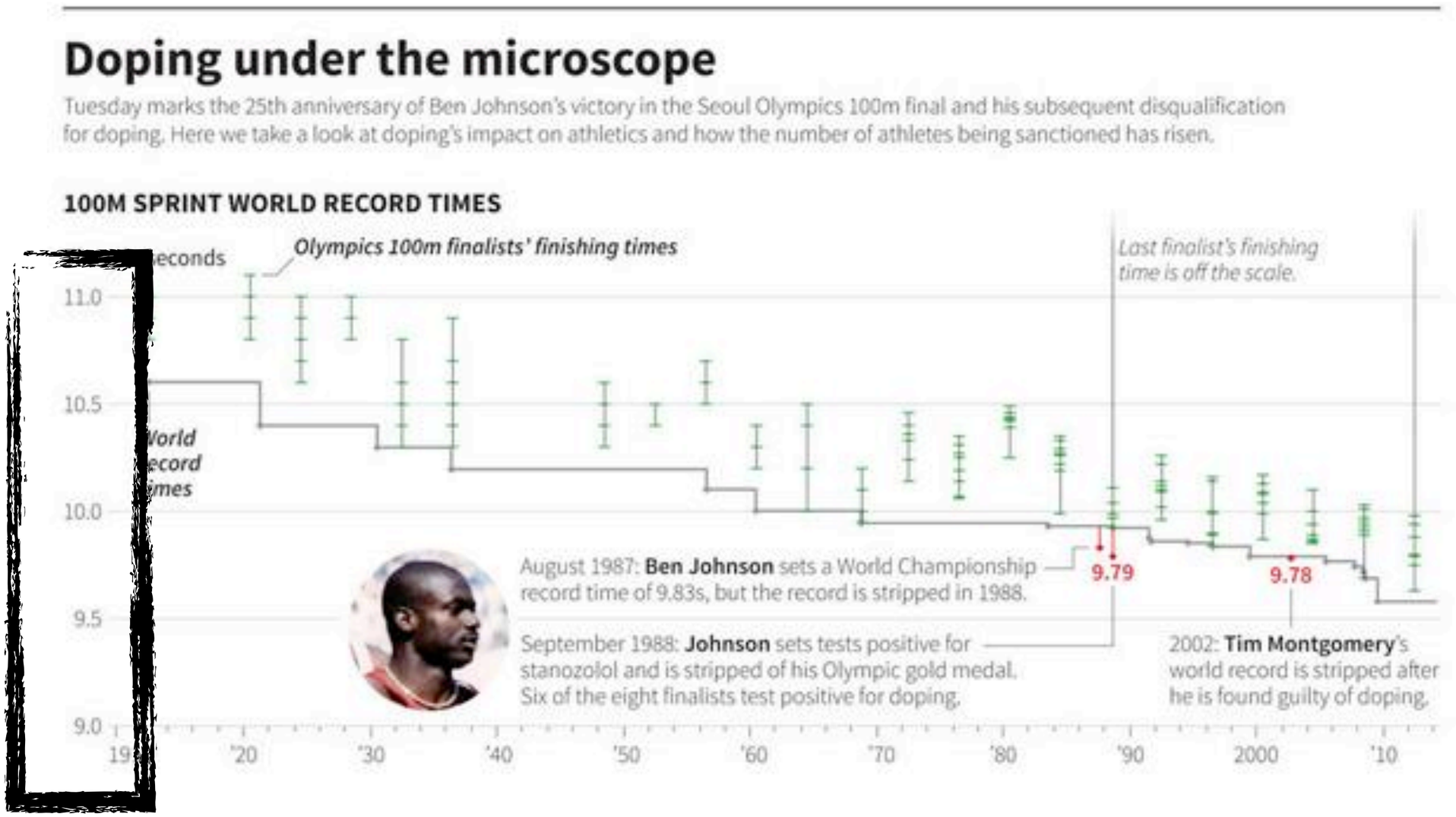
influencing factors

chart scale

barplots should never have truncated axis

lineplots *may* have them, since size is not used for encoding

choose a meaningful y-axis, keeping the trustworthiness of the graph



Stage 4

Developing your design **solution**

Stage 3

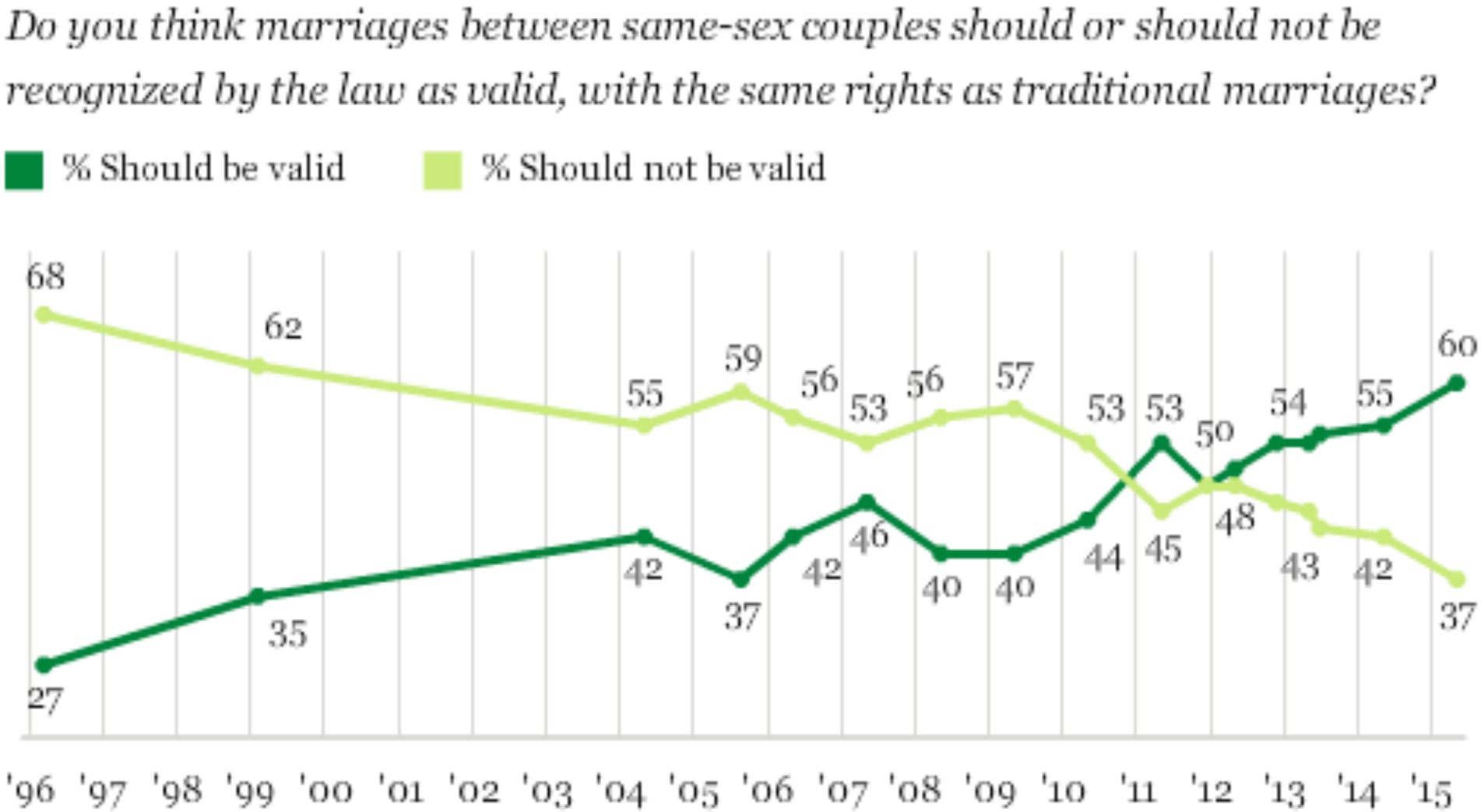
Establishing your **editorial** thinking

composition

influencing factors

chart scale

barplots should never have truncated axis
lineplots *may* have them, since size is not used for encoding



here the truncation is not correct — the sense of comparable scale is compromised

Stage 4

Developing your design **solution**

Stage 3

Establishing your **editorial** thinking

composition

influencing factors

map projection

- every projection is distorted
- the larger the area, the greater the distortion
- no projection can accommodate all map purposes
- choose damage limitation as the driving principle
- ‘equal area’ projections better for thematic mapping: distortion on shape rather than on size, thus values per region are correct
- scope of view, distance from equator, focus on land/sea can drive the choice