CASA0029: Urban Data Visualisation

Charts and Statistical Visualisations

Financial Times Visual Vocabulary:
http://ft.com/vocabulary

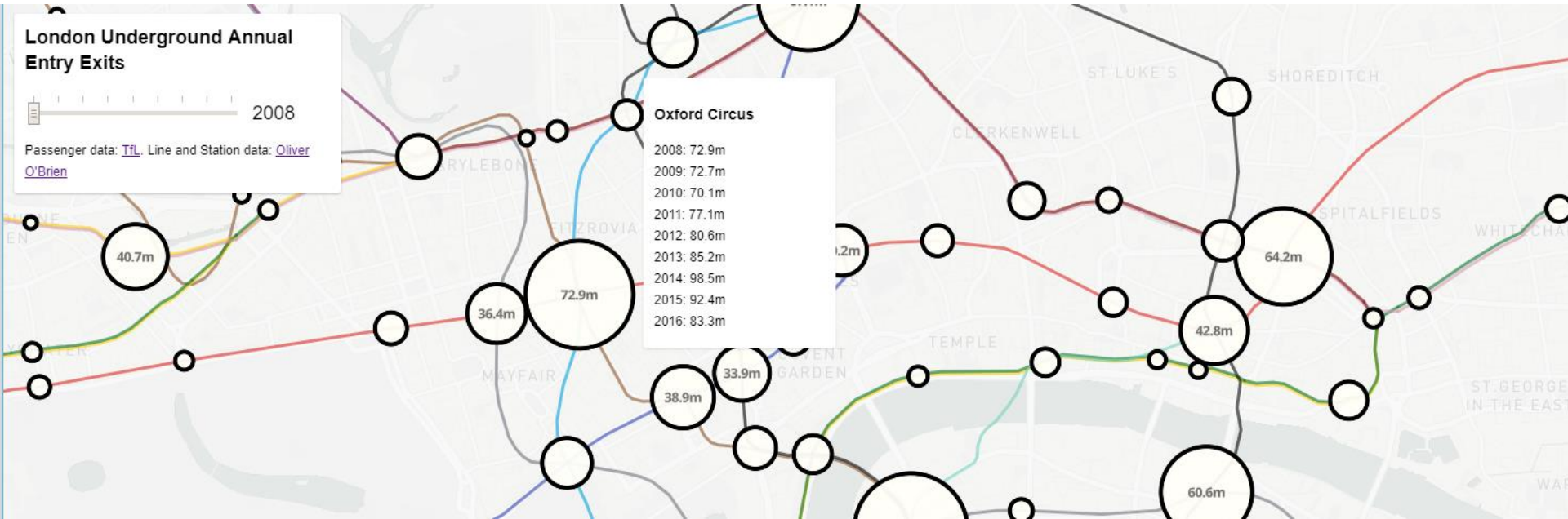
Duncan A Smith

Centre for Advanced Spatial Analysis, UCL

Revision - Practical from Last Week…

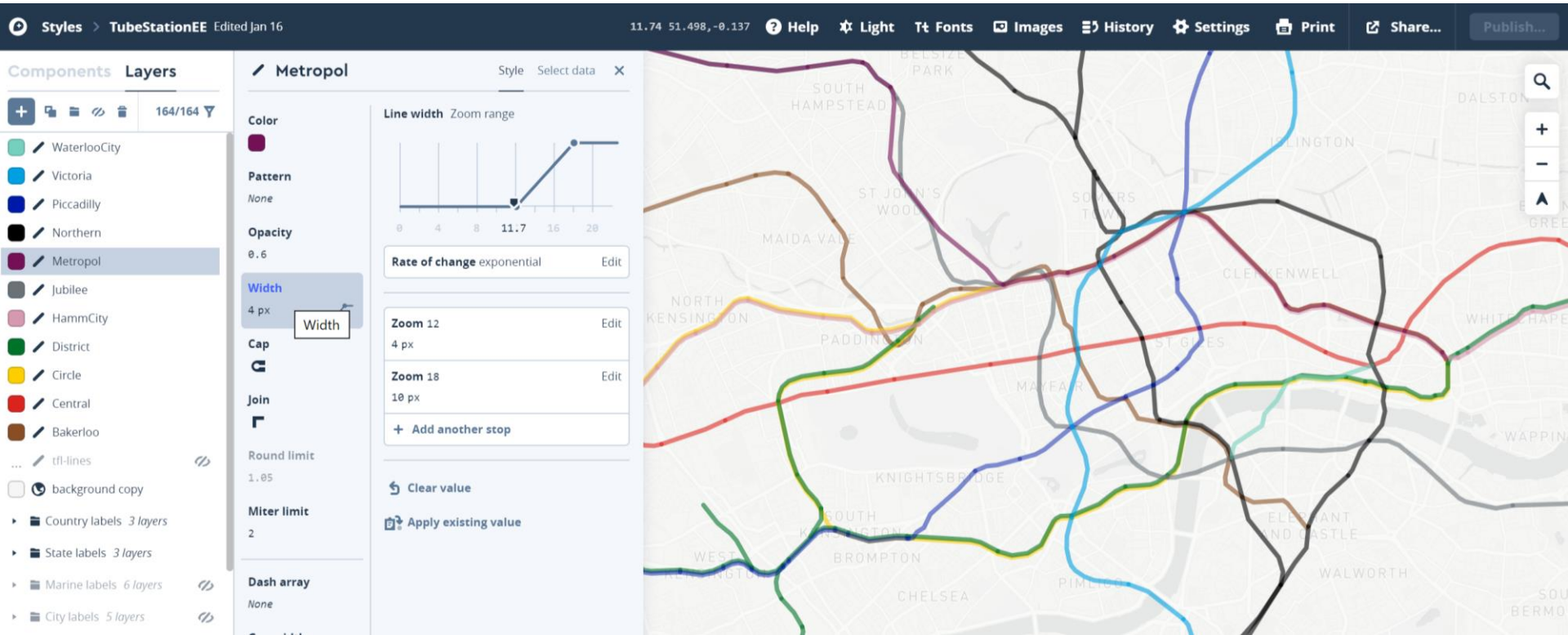
Studio for Fixed Background Layers; Mapbox.gl JS for Dynamic Interactive Layers
Mapbox example combined a background Style of Tube Lines (created in Studio) with a Tilelayer of Tube Stations as proportional circles, added via Mapbox GL JavaScript.

The background style is fixed cartographically, while the Tilelayer can be changed dynamically in the browser by changing the 'paint properties' of the layer. In this case the radius of the circles changes (and the station labels) after user input.



Mapbox Studio Basemap Design

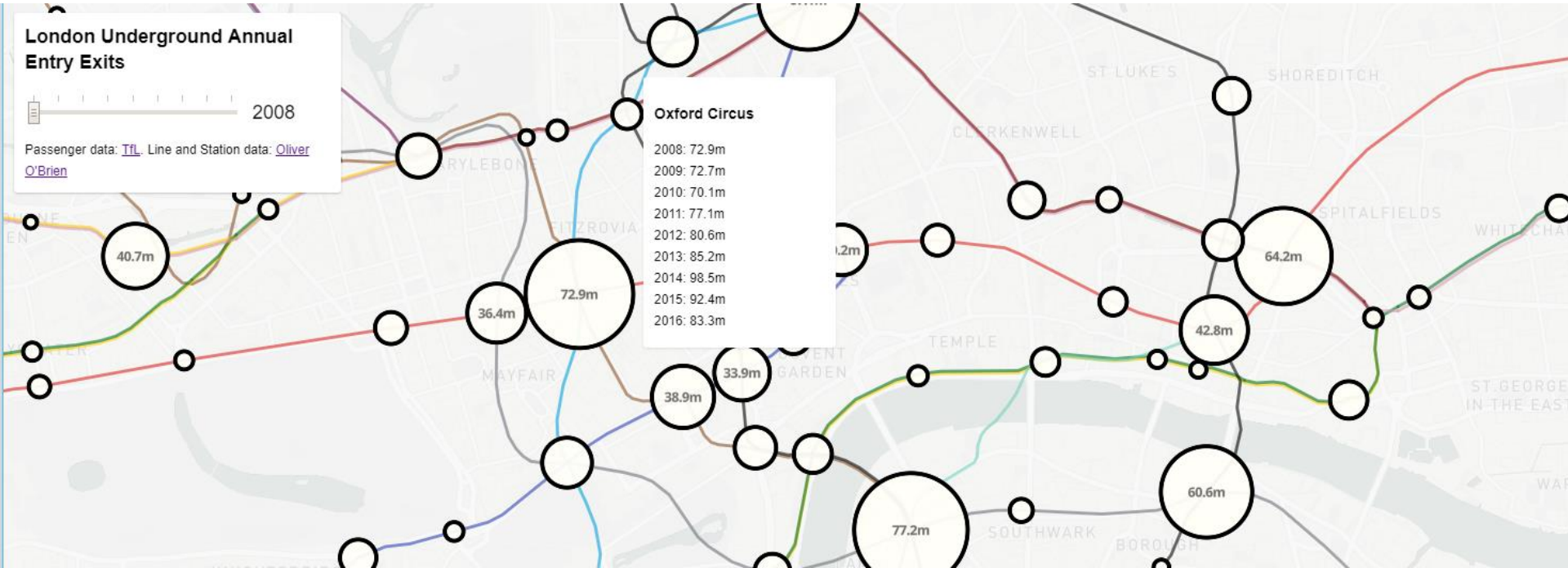
Mapbox Studio offers a high degree of cartographic control. Styling by data variables and by zoom level. The final style is largely ‘fixed’ and so is generally used for base maps with data layers visualised on top of the basemap.



Events in JavaScript

An ‘event listener’ is attached to the slider input. An event listener function runs when a particular specified user interaction occurs (e.g. mouseover, click…). In this case, the function ‘setYear(year)’ is run when the user changes the input slider.

The function setYear gets the Paint Property ‘circle-radius’ from the TubeStation layer. It changes the column used for the radius to the new column year the user has selected, and then sets the paint property to this new column.



Common User Events in Interactive Mapping

User events triggering changes to the properties of visualisations very common and used in lots of interactive visualisation-

Highlighting Features

Mouseover (hover) event triggers the visibility of the highlight layer. You need the ID of the feature highlighted for this to work.

Swapping Layers

One layer removed and another added. Usually a radio-button type interface is used for this kind of interaction.

Changing Visibility of Features

Altering the visibility or alpha value of a layer. Could be checkbox for visible/invisible; or slider for incremental control of transparency.

Filtering Features

Similar to GIS, attribute and spatial filtering possible. Could be based on selecting categories in the legend; or mouseover / selection interaction on the map.

Zooming to a Location

Could use a searchbox, or links to specific popular locations.

Today's Lecture Overview

1. **Data Visualisation Principles**
2. **Choosing Chart Types**
3. **Charts, Maps and Animation**

Practical-

4. Creating Charts in JavaScript

Overview

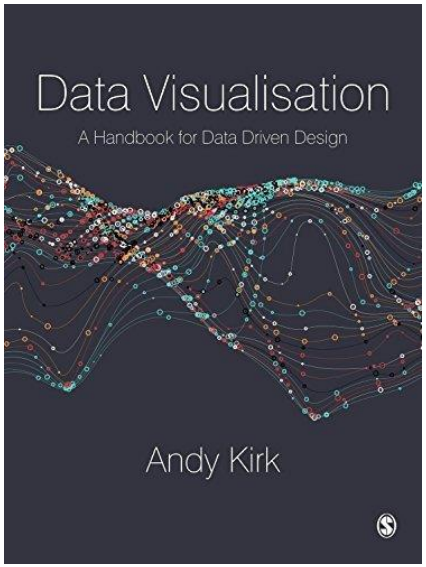
1. **Data Visualisation Principles**
2. Choosing Chart Types
3. Charts, Maps and Animation

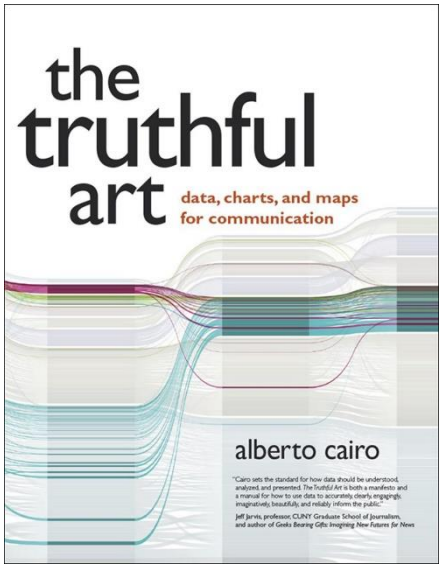
Practical-

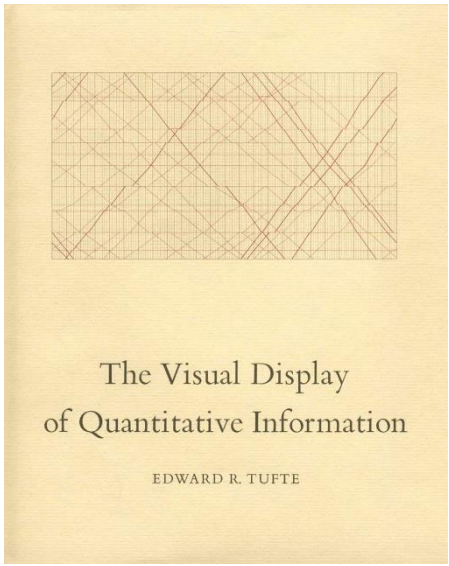
5. Creating Charts in JavaScript

Recommended Reading on Data Visualisation

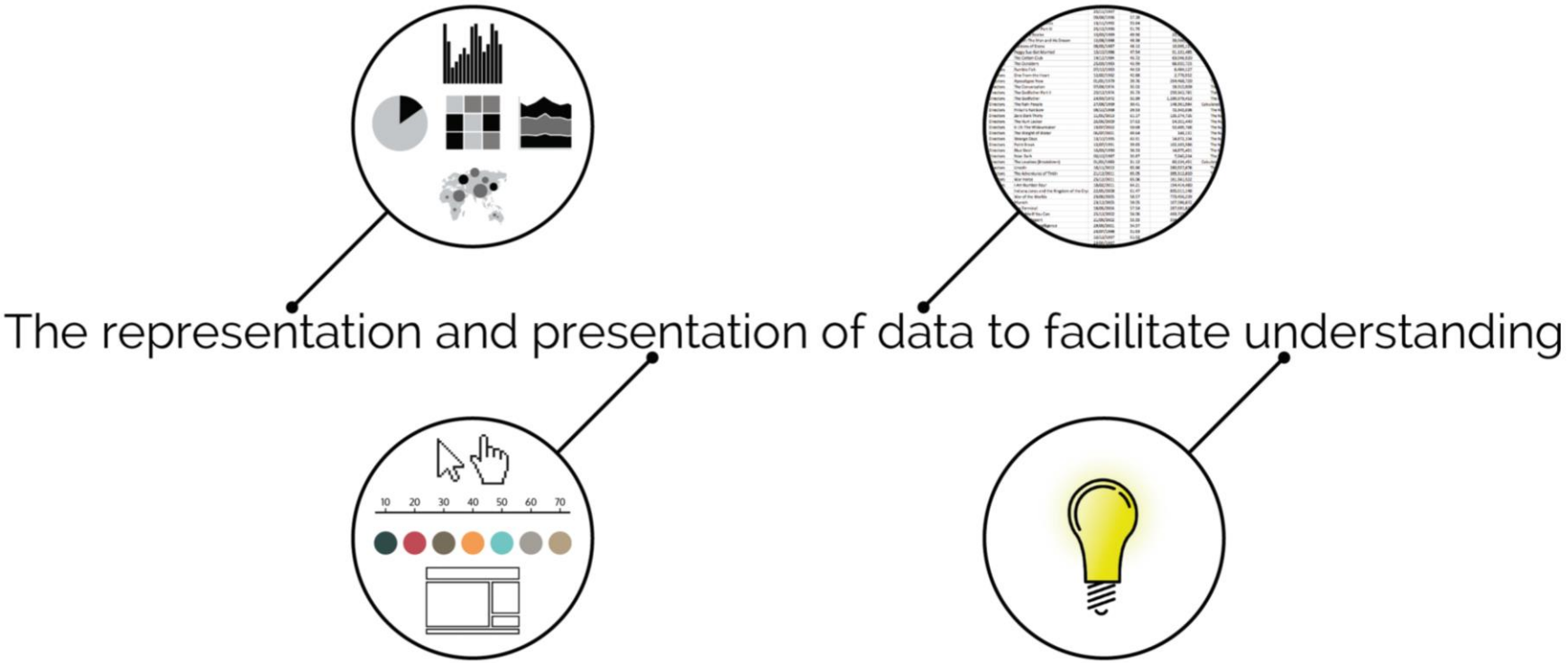
Good guides to data visualisation principles- Andy Kirk, Alberto Cairo and Edward Tufte. Recommended reading-



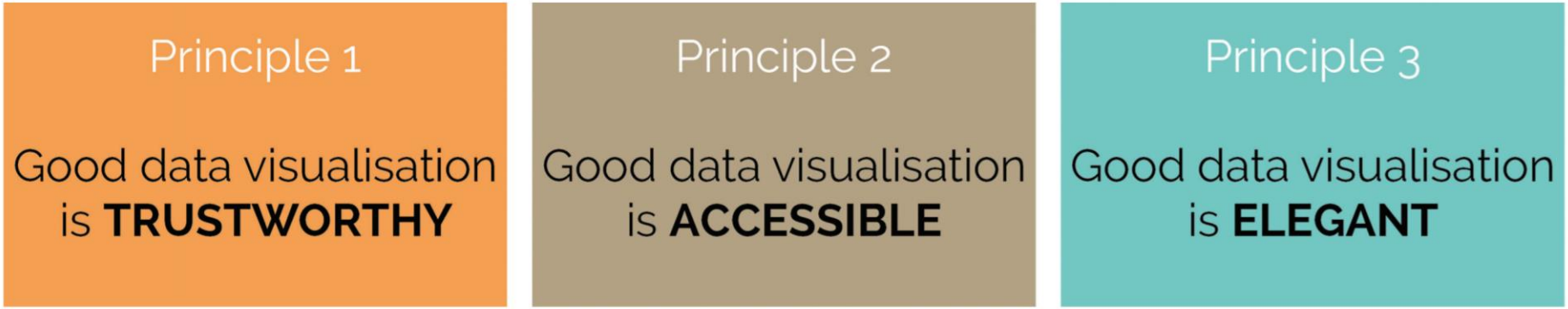




Definition of Data Visualisation (Kirk, 2016)-

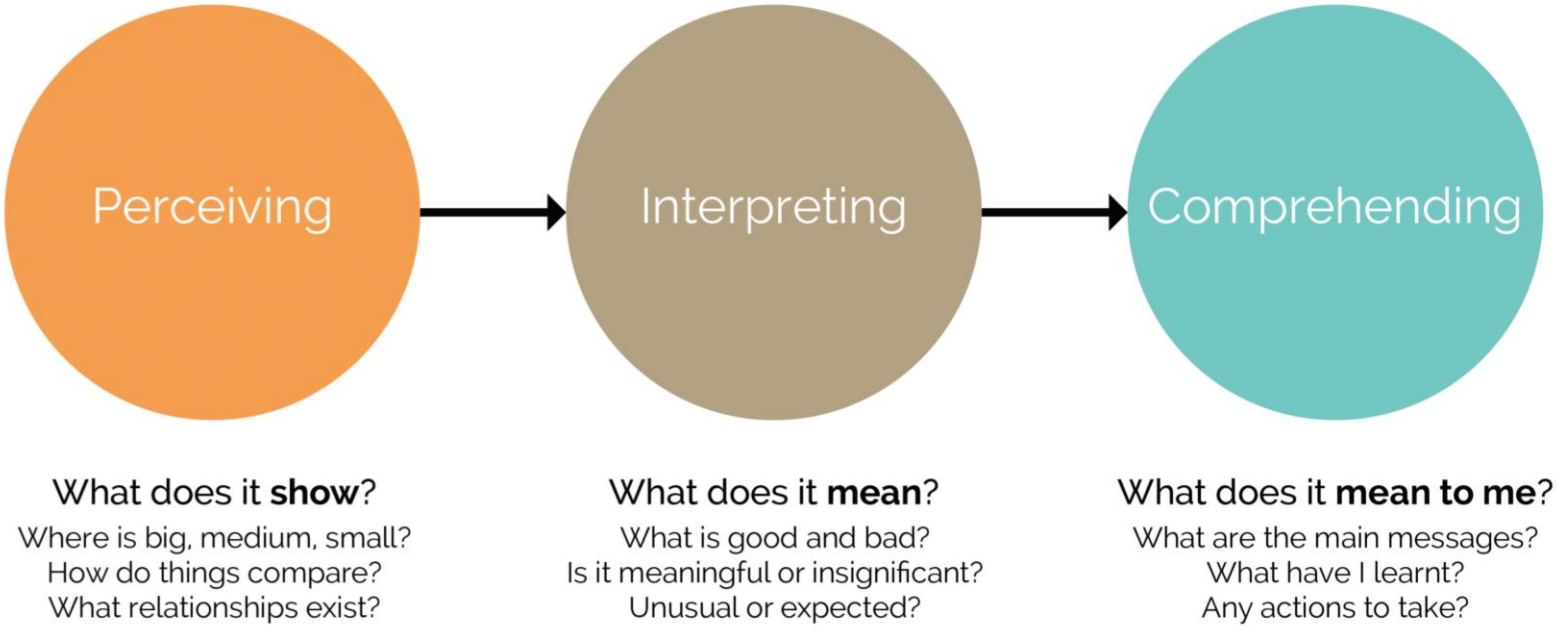


Three Principles of Good Visualisation Design (Kirk, 2016)-



For research applications, could add ‘insightful’ as forth principle

Three stages of understanding in relation to data visualisation (Kirk, 2016)-

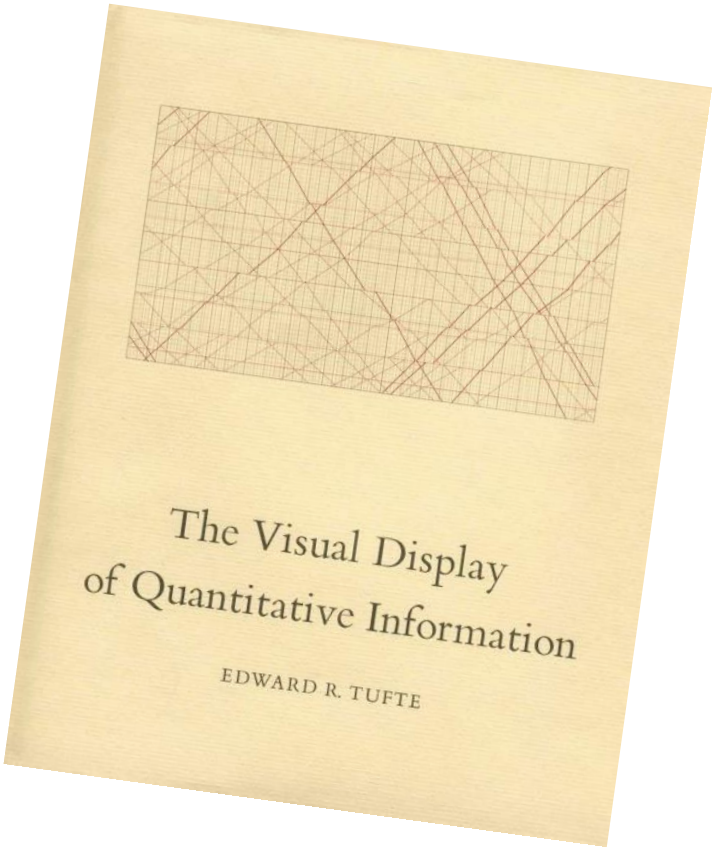


Clarity, Precision & Efficiency- Tufte

Foundational figure in data visualisation- Edward Tufte. Classic work “The Visual Display of Quantitative Information” (1983).

Tufte’s principles of good statistical graphics- “clarity, precision, efficiency”. Minimalist approach; recommends-

- Show the data;
- Induce the viewer to think about the substance rather than the methodology or graphic design;
- Avoid distorting the data;
- Make large datasets coherent;
- Reveal the dataset at several levels of detail;
- Integrate the statistical and verbal descriptions of the dataset;



Current Receipts of Government as a Percentage of Gross Domestic Product, 1970 and 1979

The slope graph illustrates Tufte's approach. Designed as a minimalist way to show 2-point time-series. All the 'ink' is 'data ink', easy to understand.

Design and Over-simplification

Abstraction central to design and visualisation. Often reduce and remove detail to avoiding ‘data-overload’. But risk of too much simplification in research applications.

Cairo (2012) argues for retaining data volume, and using design techniques to improve the legibility of large datasets for the viewer. Can utilise the visual hierarchy approach we discussed last week.

Line Chart with lots of data, design failure

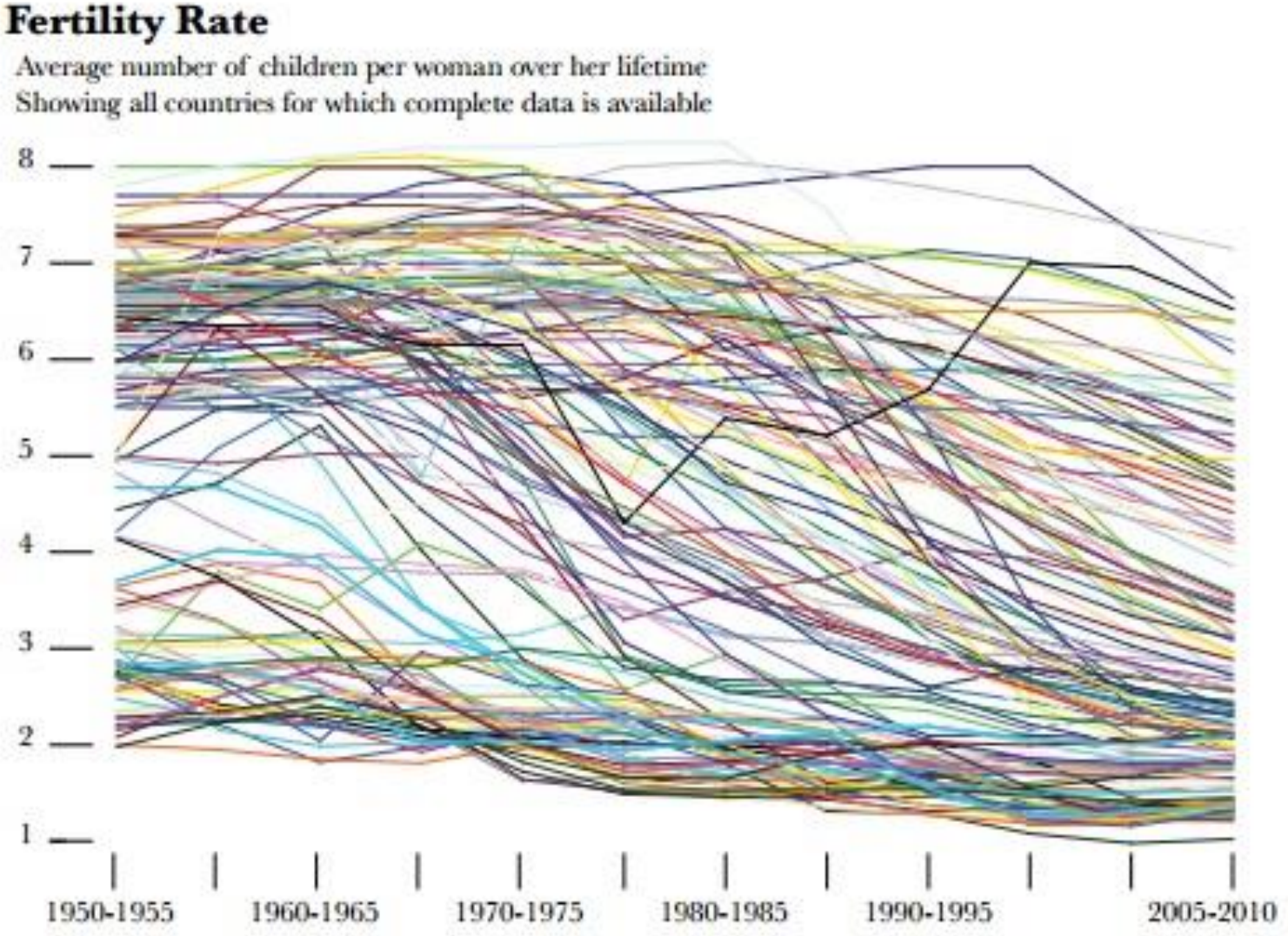
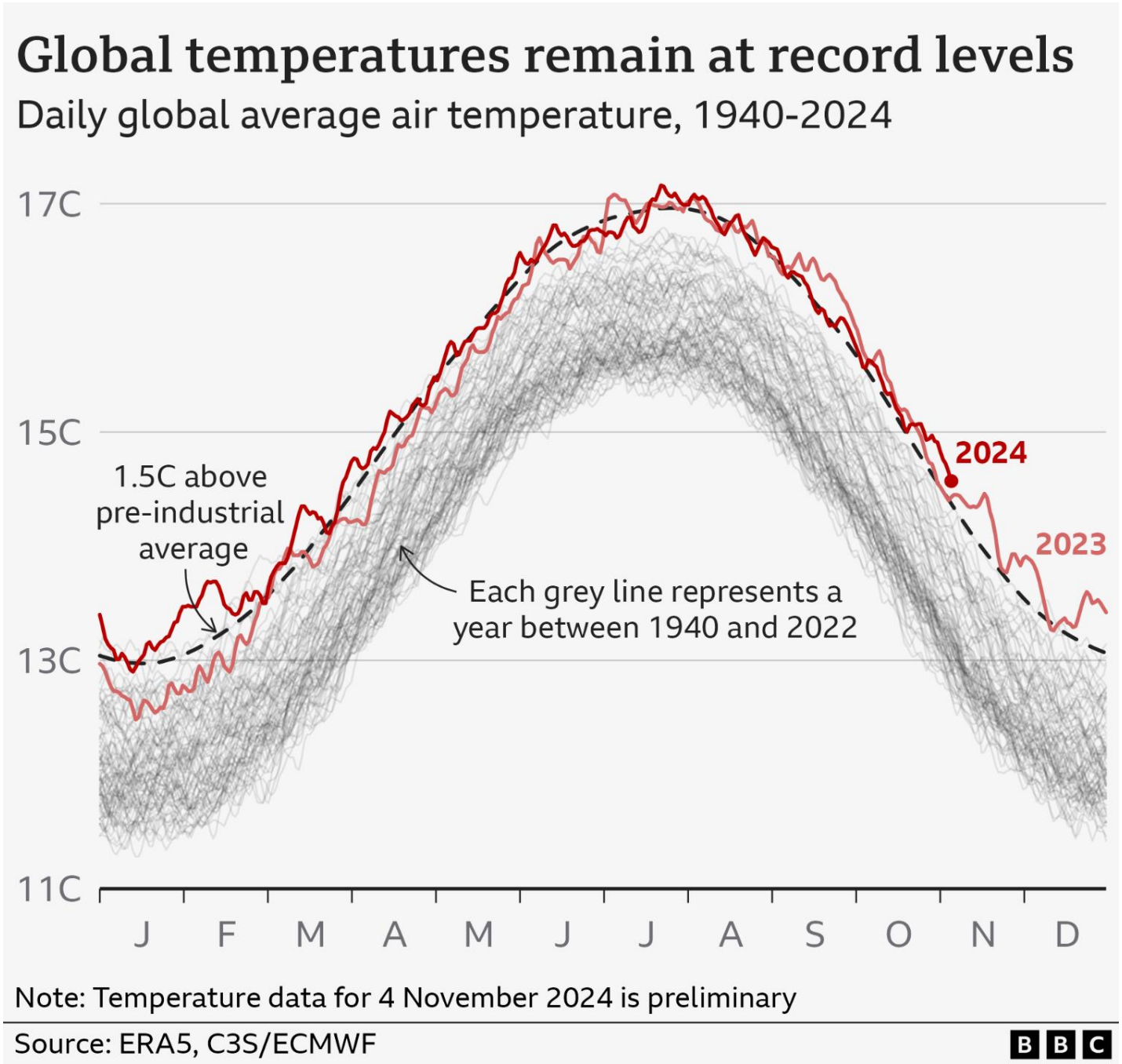


Figure 1.5 Too many lines obscure the message.

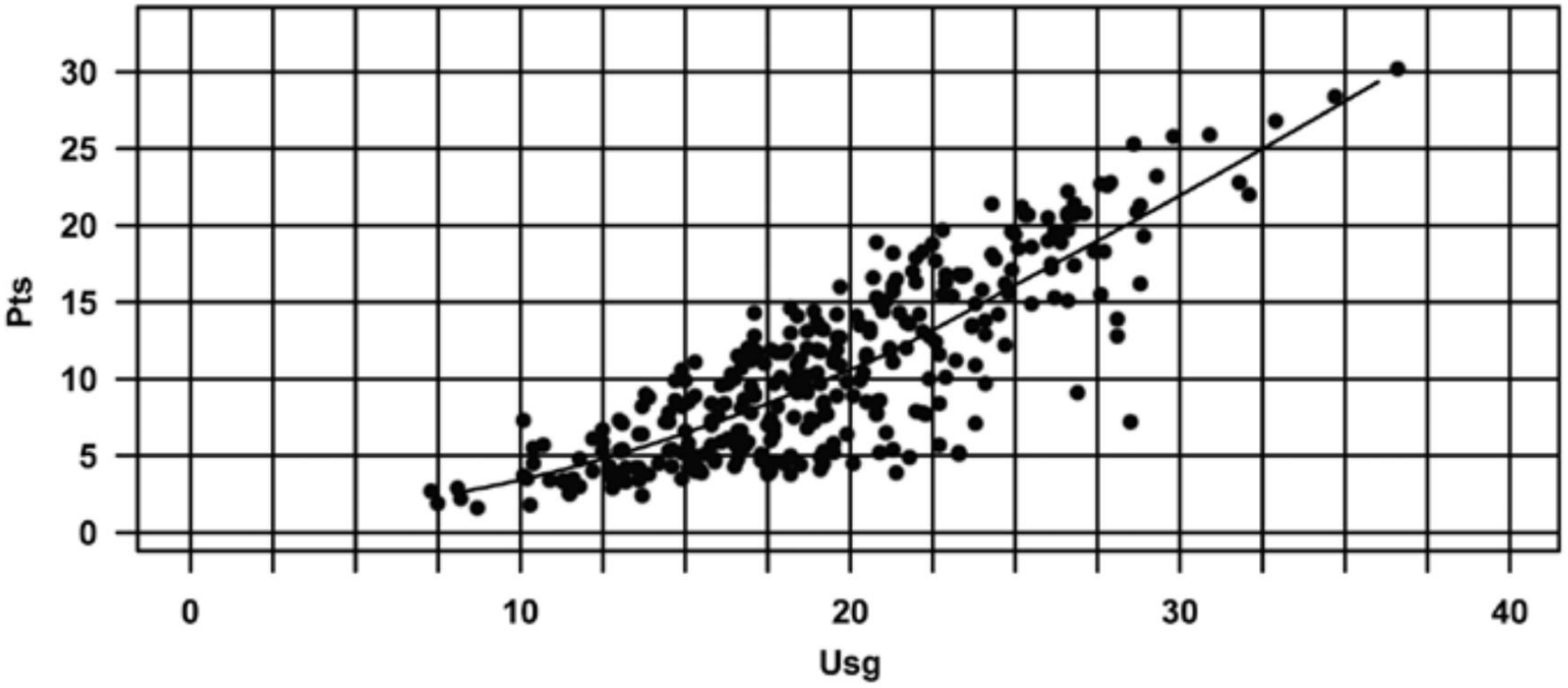
Line Chart with lots of data and visual hierarchy



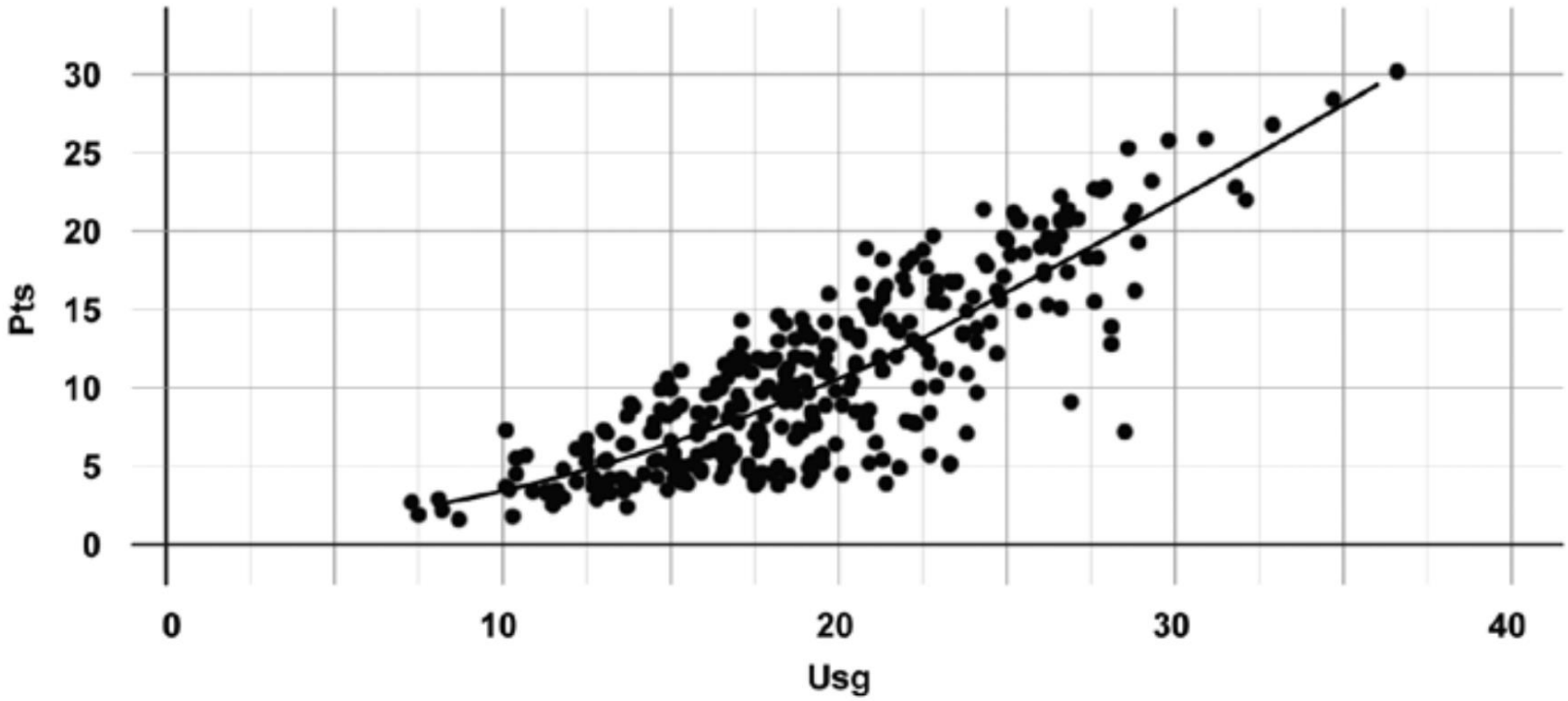
Visual Hierarchy and Chart Elements

Can also use visual hierarchy to design chart elements.

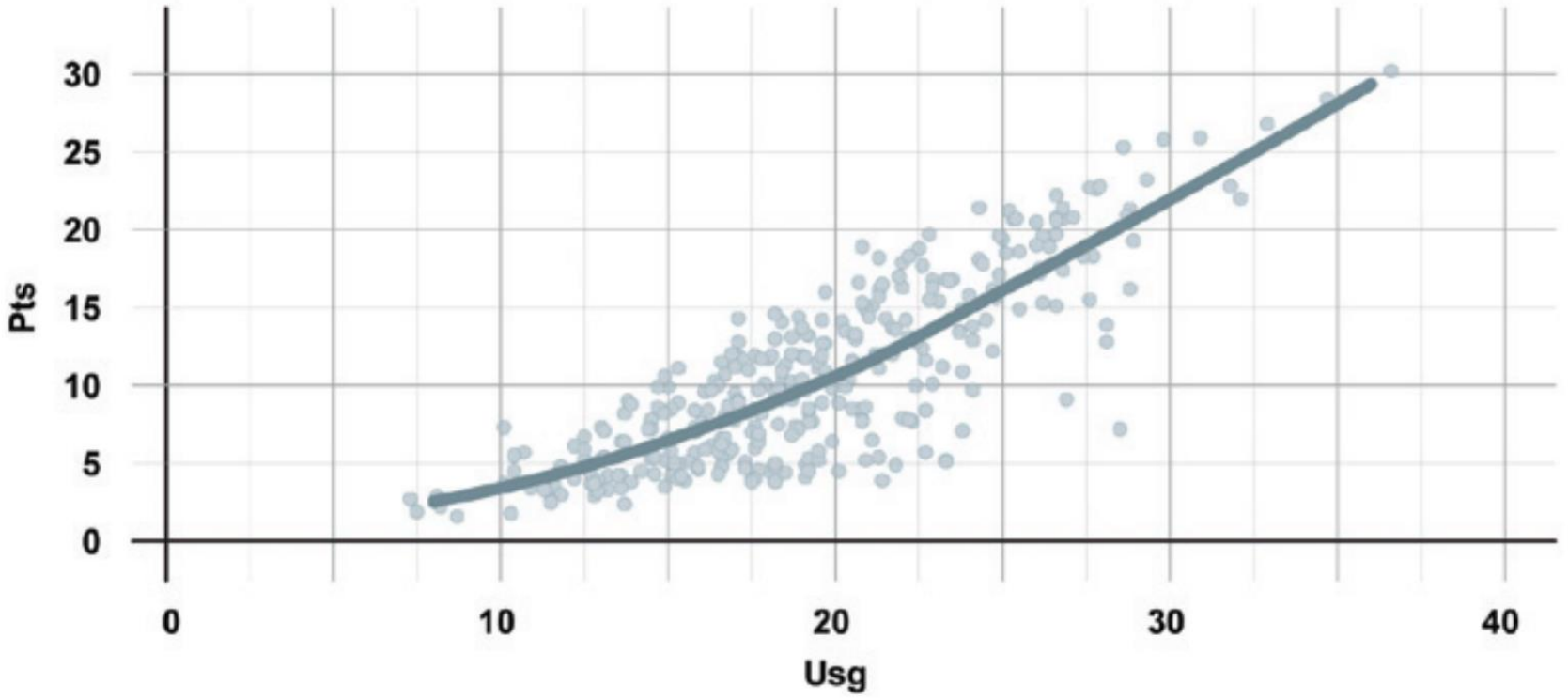
Original chart with no visual hierarchy-
(example from Nathan Yau, Data Points)



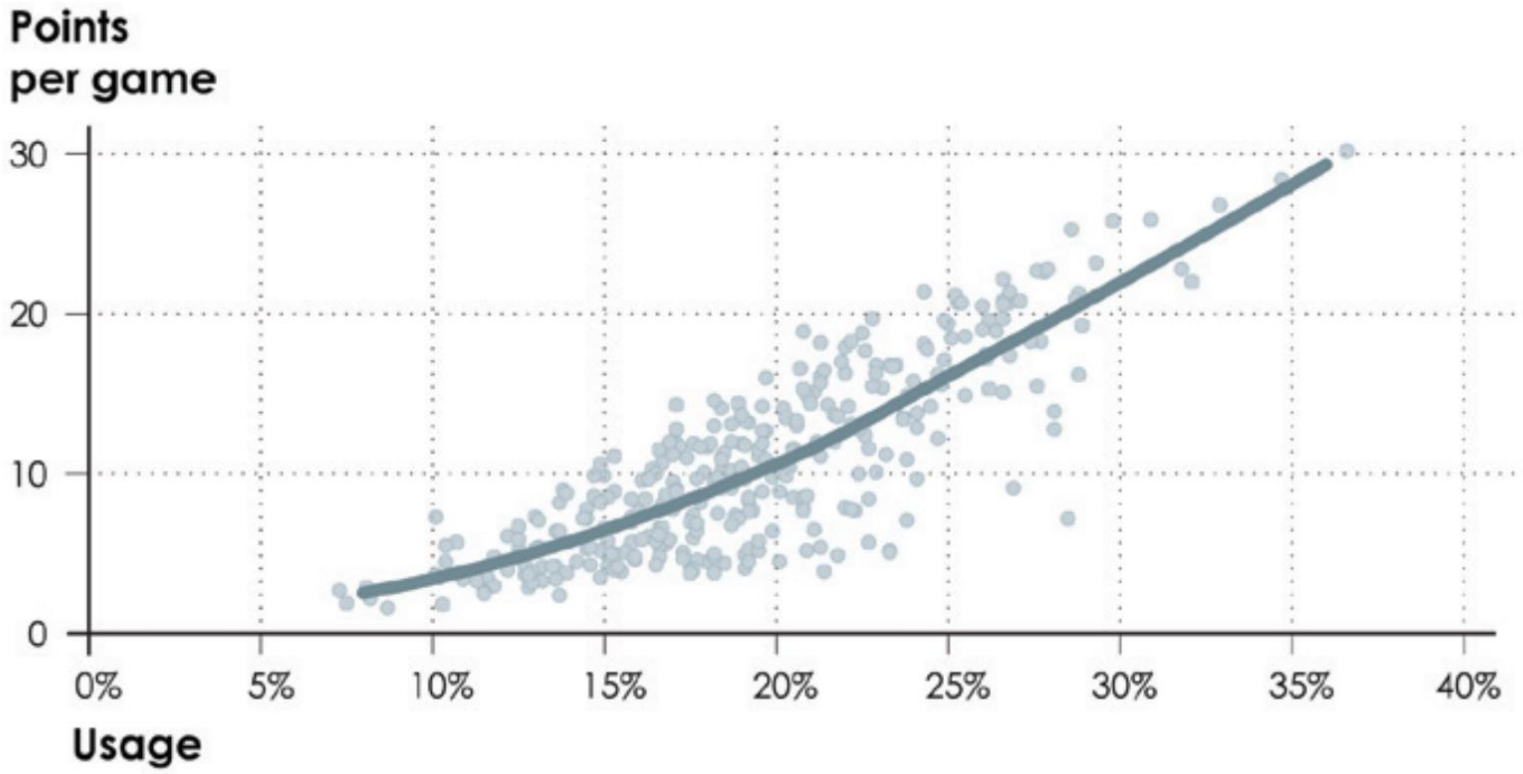
Background grid is less important, lowest in visual hierarchy. Remove emphasis-



Use colour and thickness to put trend line as most important element-



Make labels easy to understand and clear.
Final chart is a significant improvement-



1. Data Visualisation Principles
2. **Choosing Chart Types**
3. Charts, Maps and Animation

Practical-

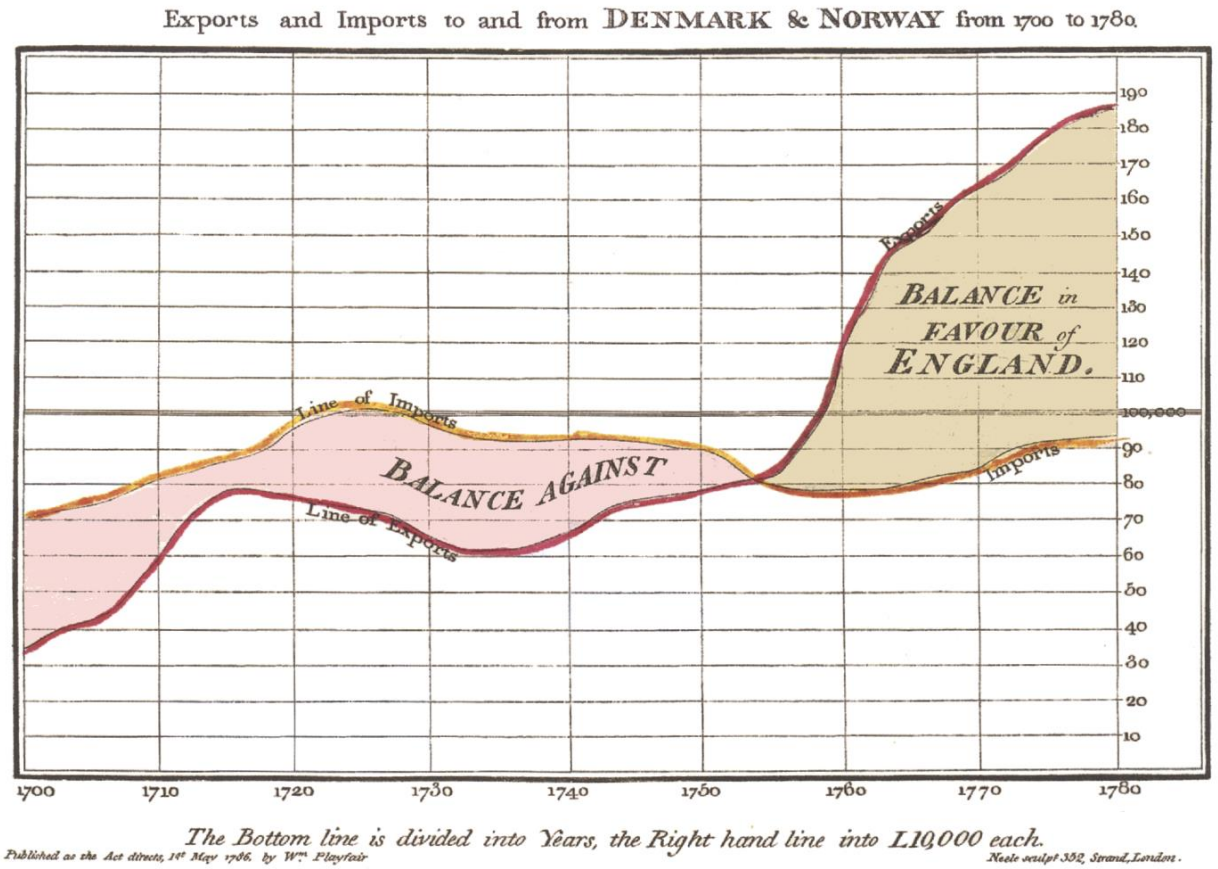
5. Creating Charts in JavaScript

Inventing Charts in the 18th Century

Line charts designed as efficient way to show time-series data.

Playfair, 18th century Scottish economist, key figure in early chart design. Argues charts more engaging and memorable than tables of data-

"On inspecting any one of these Charts attentively, a sufficiently distinct impression will be made, to remain unimpaired for a considerable time, and the idea which does remain will be simple and complete"
(The Commercial and Political Atlas, 1786)



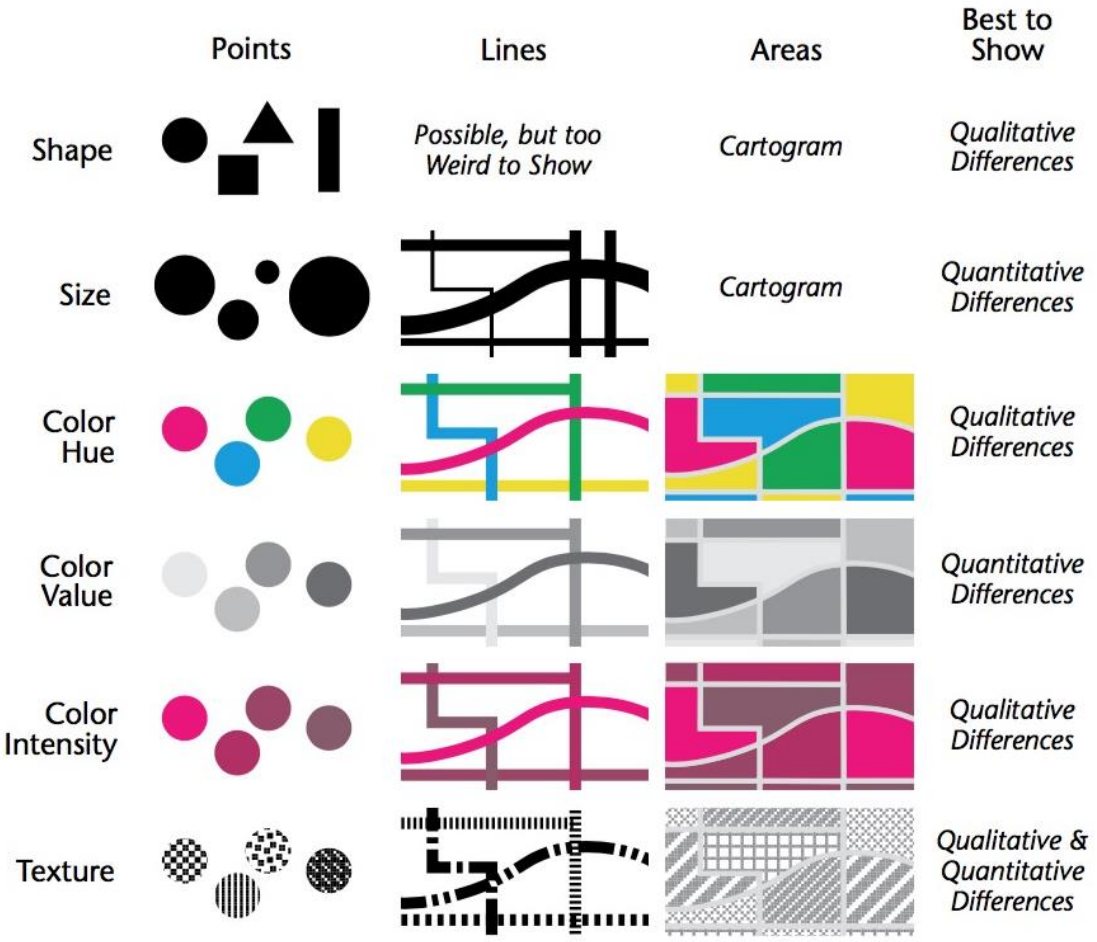
Choosing Chart Types

Chart Types Developed for Particular Visual Tasks
Bar/column charts quickly enable magnitude comparisons; line charts designed for time-series data.

Simple or Advanced Charts?
Simple charts often most efficient and legible for a larger audience. Partly due to familiarity; also cognitive reasons for efficiency of classic chart types.

There are cases where more advanced chart types can really add to engagement and data exploration. Good to experiment with more ambitious chart types in this module.

Visual Channels in Mapping Similar to Chart Design



Krygier and Wood (2016), Making Maps. Influenced by Semiologie Graphique, Jacques Bertin (1983). Need to consistently use visual channels, plan which channel most appropriate for data type.

Visual Channels and Cognitive Efficiency

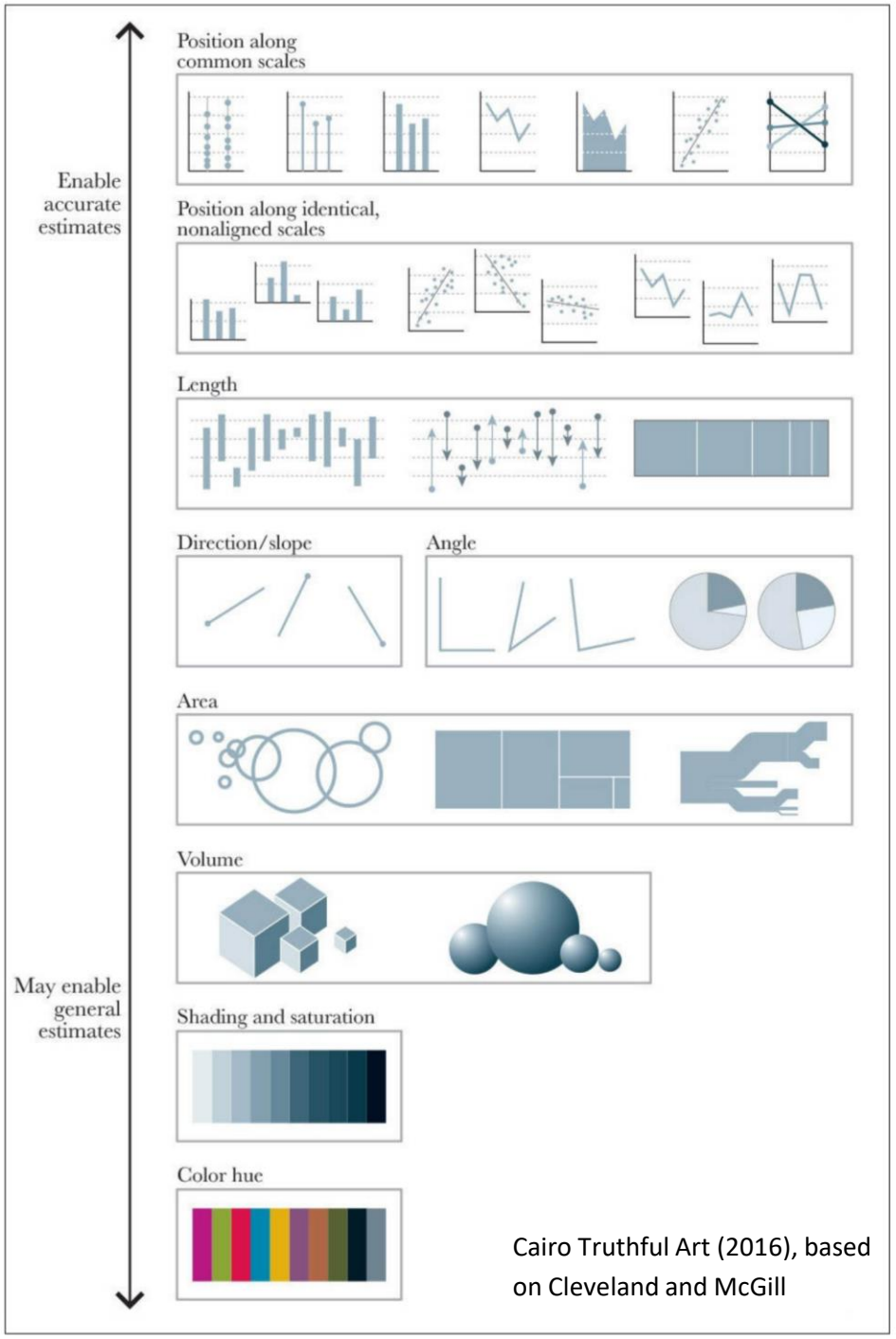
Length/Position on Common Scale Most Efficient

Human visual system strong at making comparisons across a common scale. On a bar/column chart or line chart can easily identify ranks and differences in magnitude (e.g. x is twice as big as y).

Angle, Area and Colour Weaker Channels for Magnitude

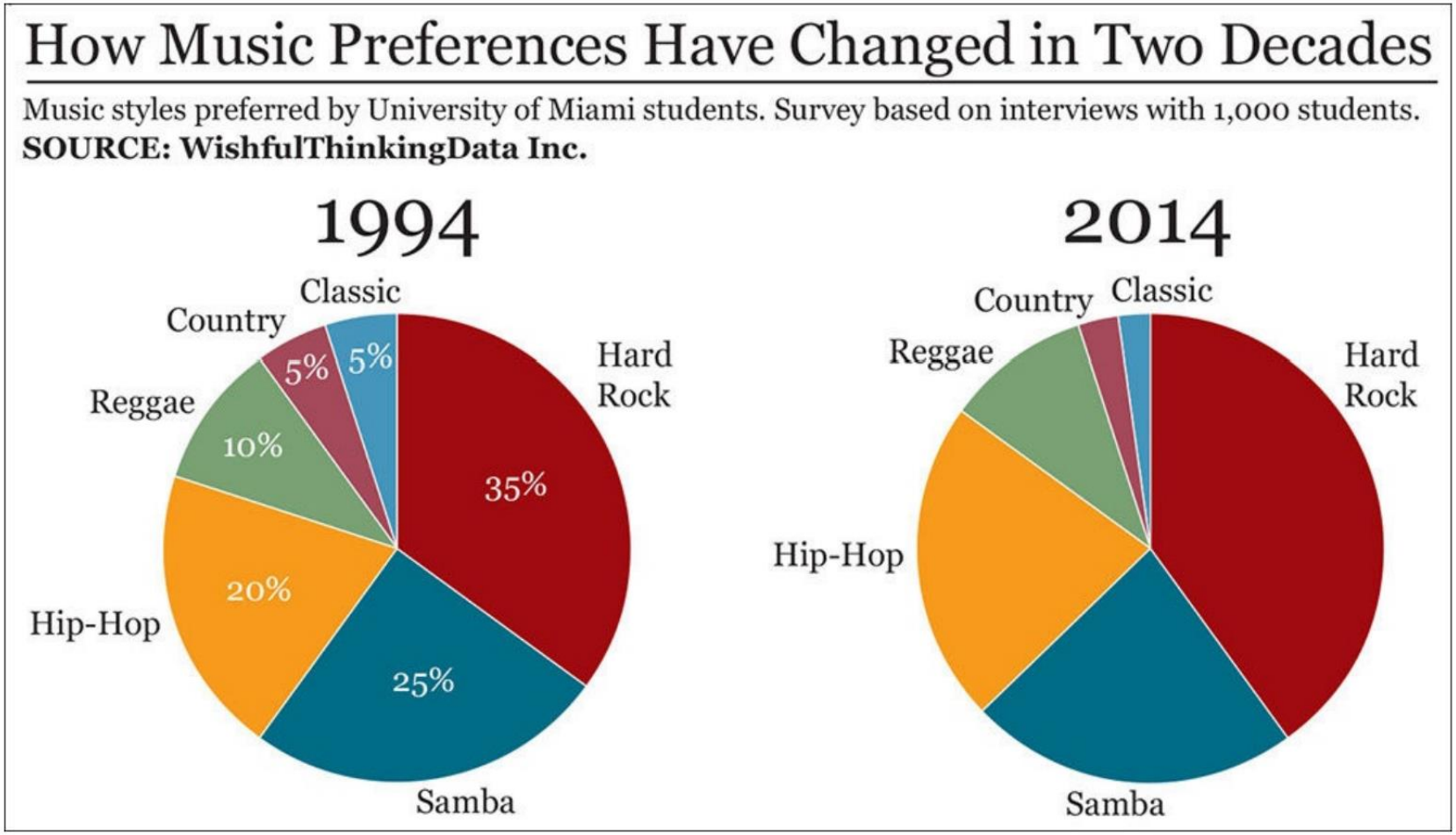
Pie charts use angle, and are not set on a common scale. Comparisons more difficult to make and poor for change over time.

Area charts allow rank comparisons, but are weaker for precise differences in magnitude. Colour/saturation used in maps are even weaker for magnitude comparisons.



Cairo Truthful Art (2016), based on Cleveland and McGill

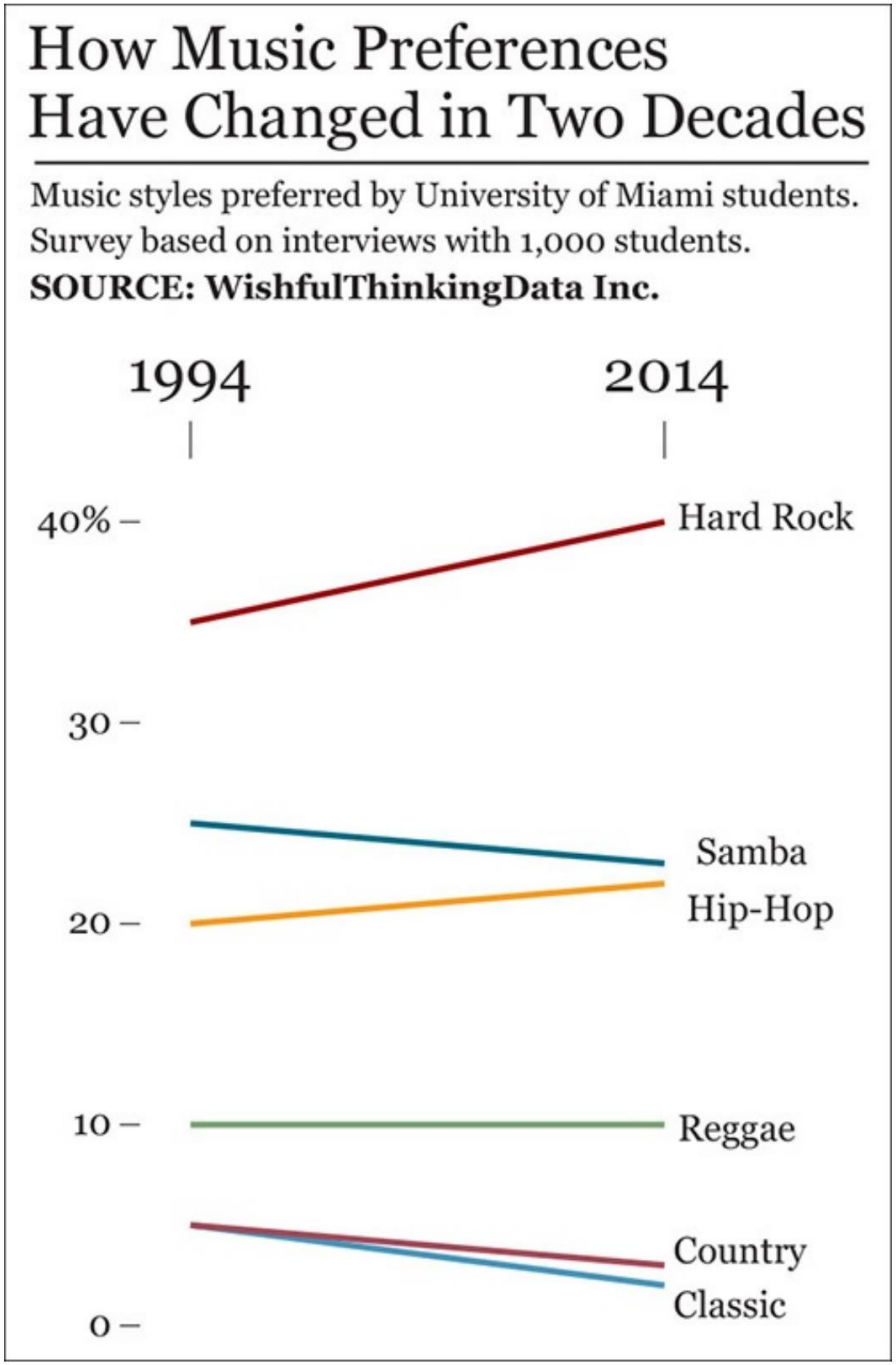
Pie-charts based on angle on a non-aligned scale. Users cannot make accurate magnitude assessments. Poor for understanding change over time. In example below, has Samba category grown or shrunk? How much has Hip-Hop grown by?



Cairo (2016) Truthful Art

If the same data is visualised as a slope chart, this allows changes over time to be assessed very easily.

Slope chart based on position on a common scale.



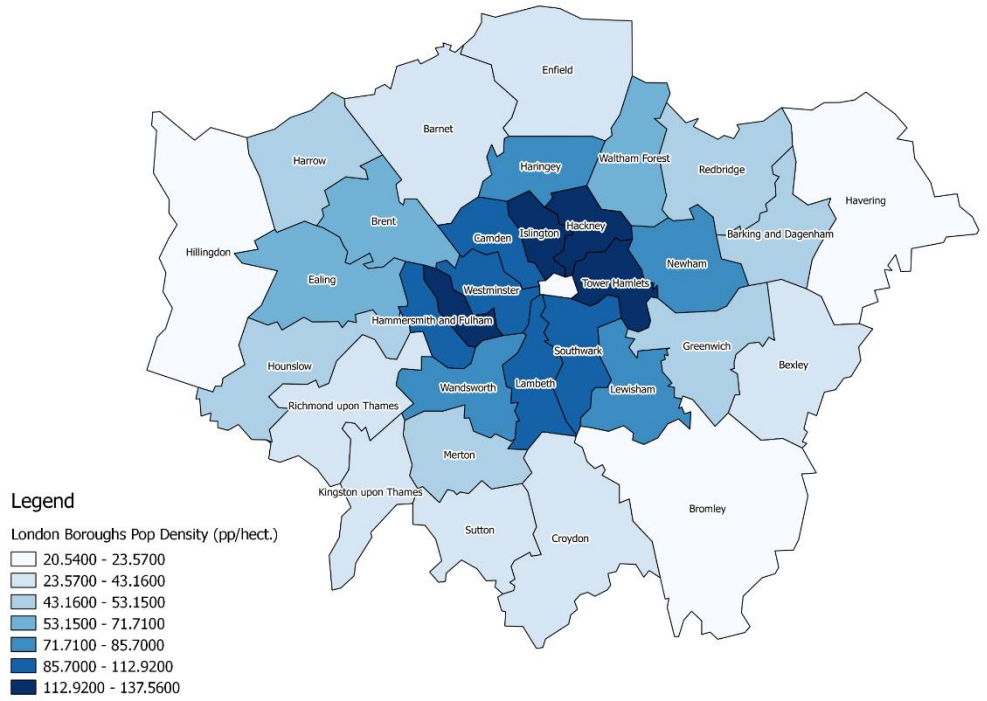
Cairo (2016) Truthful Art

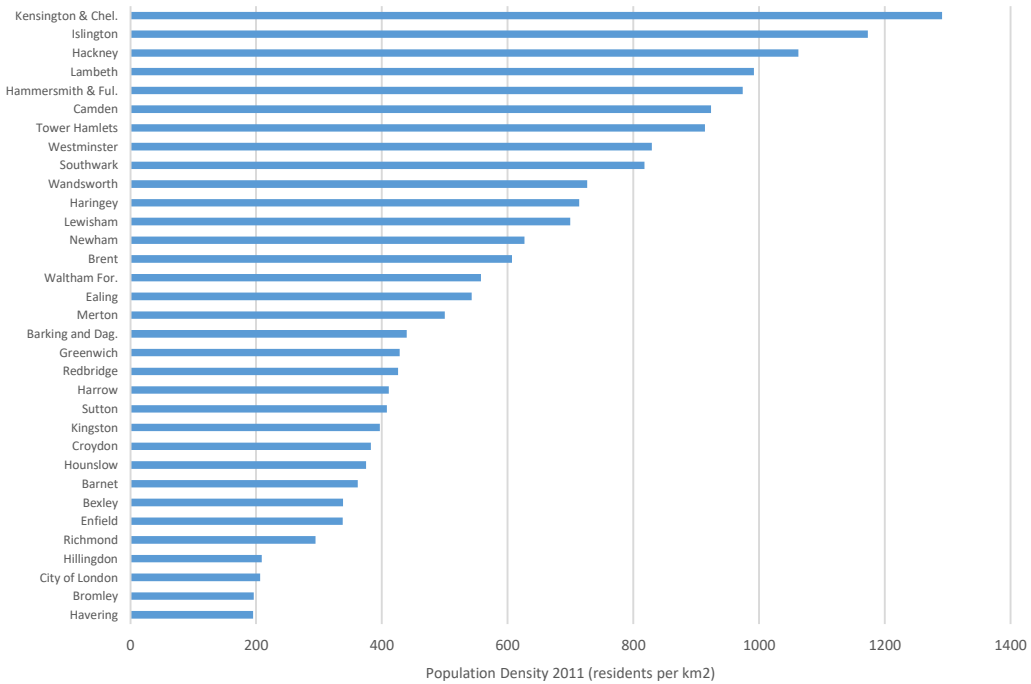
Charts versus Maps

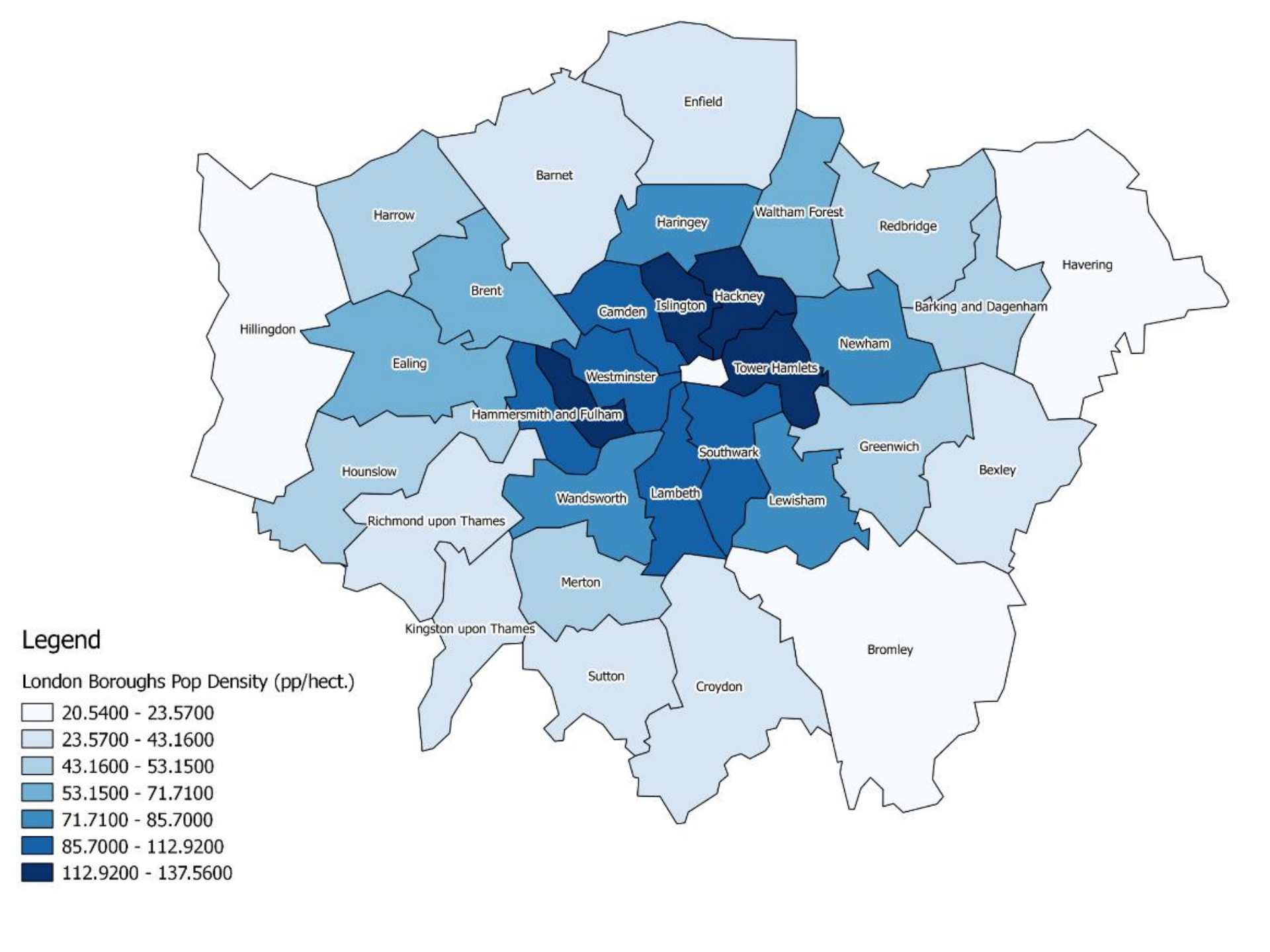
Standard choropleth maps based on colour shading. Work well for general estimates such as rank. But inefficient for magnitude.
For example, Tower Hamlets is clearly denser than Newham in the map, but how much denser? Have to refer to legend, and even then difficult to estimate.

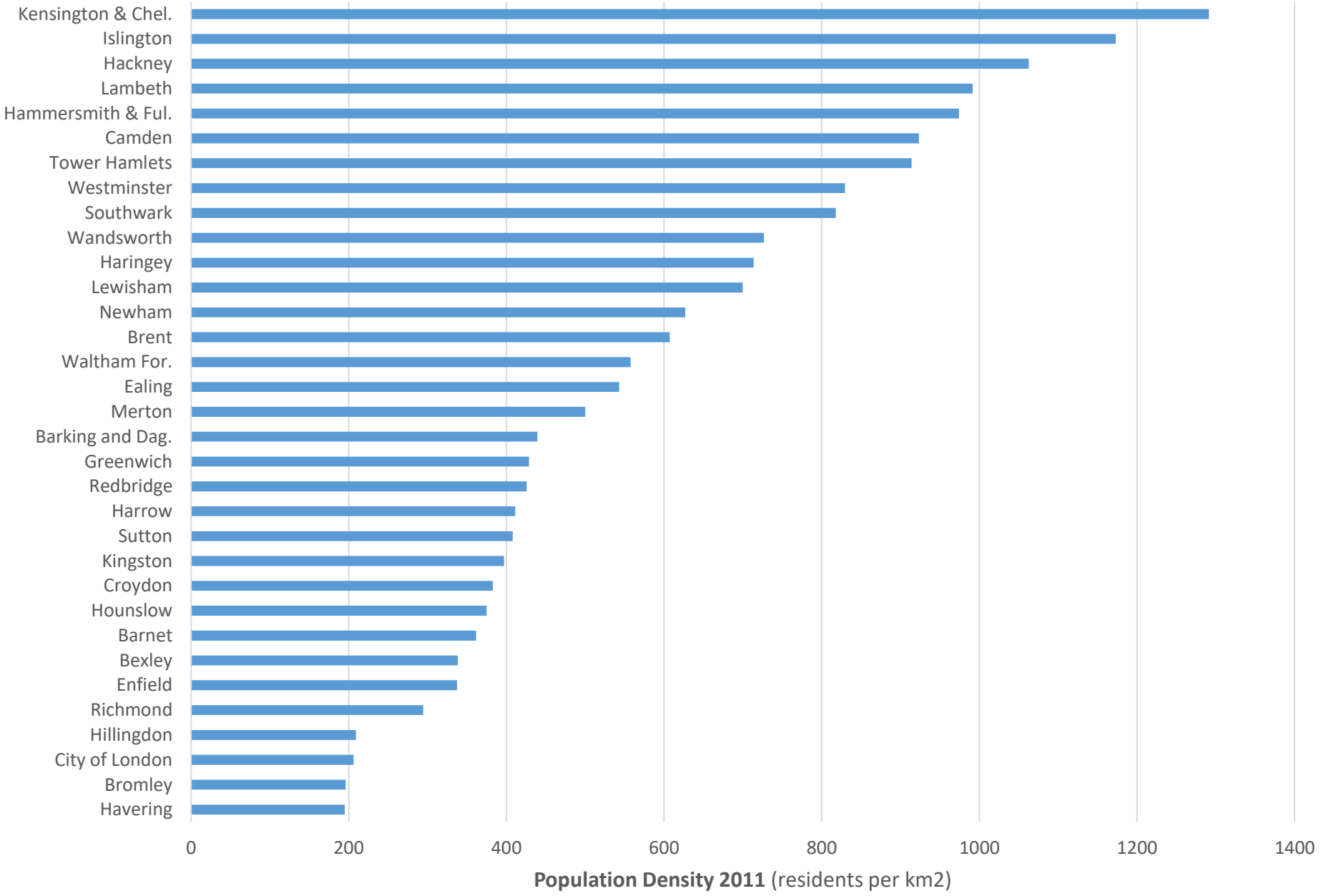
Proportional symbols possible mapping alternative using area rather than colour, but also limits to how precisely users can estimate area.

When magnitude or change over time is the main message (rather than spatial patterns), then charts generally perform better than maps. Can also consider including both.






Enfield
Barnet
Harrow
Haringey
Waltham Forest
Redbridge
Havering
Brent
Camden
Islington
Hackney
Barking and Dagenham
Hillingdon
Ealing
Newham
Westminster
Tower Hamlets
Hammersmith and Fulham
Southwark
Greenwich
Bexley
Hounslow
Wandsworth
Lambeth
Lewisham
Richmond upon Thames
Merton
Kingston upon Thames
Bromley
Sutton
Croydon
Legend
London Boroughs Pop Density (pp/hect.)
20.5400 - 23.5700
23.5700 - 43.1600
43.1600 - 53.1500
53.1500 - 71.7100
71.7100 - 85.7000
85.7000 - 112.9200
112.9200 - 137.5600


Kensington & Chel.
Islington
Hackney
Lambeth
Hammersmith & Ful.
Camden
Tower Hamlets
Westminster
Southwark
Wandsworth
Haringey
Lewisham
Newham
Brent
Waltham For.
Ealing
Merton
Barking and Dag.
Greenwich
Redbridge
Harrow
Sutton
Kingston
Croydon
Hounslow
Barnet
Bexley
Enfield
Richmond
Hillingdon
City of London
Bromley
Havering
0
200
400
600
800
1000
1200
1400
Population Density 2011 (residents per km2)

Choosing Chart Types

Lots of guides to help you choose between chart types. Kirk and Cairo books both good overviews of chart design.

Good online guide is the Financial Times Visual Vocabulary (github resource includes templates)- http://ft.com/vocabulary

Visual vocabulary

Designing with data

There are so many ways to visualise data - how do we know which one to pick? Use the categories across the top to decide which data relationship is most important in your story, then look at the different types of chart within the category to form some initial ideas about what might work best. This list is not meant to be exhaustive, nor a wizard, but is a useful starting point for making informative and meaningful data visualisations.

FT graphic: Alan Smith, Chris Campbell, Ian Bott, Liz Faunce, Graham Parrish, Billy Ehrenberg-Shannon, Paul McCallum, Martin Stabe
Inspired by the Graphic Continuum by Jon Schwabish and Severino Ribecca

ft.com/vocabulary

Deviation

Emphasise variations (+/-) from a fixed reference point. Typically the reference point is zero but it can also be a target or a long-term average. Can also be used to show sentiment (positive/neutral/negative).

Example FT uses
Trade surplus/deficit; climate change

Diverging bar

A simple standard bar chart that can handle both negative and positive magnitude values.

Diverging stacked bar

Perfect for presenting survey results which involve sentiment (eg disagree/neutral/agree).

Spine

Splits a single value into two contrasting components (eg male/female).

Surplus/deficit filled line

The shaded area of these charts allows a balance to be shown – either against a baseline or between two series.

Correlation

Show the relationship between two or more variables. Be mindful that, unless you tell them otherwise, many readers will assume the relationships you show them to be causal (i.e. one causes the other).

Example FT uses
Inflation and unemployment, income and life expectancy

Scatterplot

The standard way to show the relationship between two continuous variables, each of which has its own axis.

Column + line timeline

A good way of showing the relationship between an amount (columns) and a rate (line).

Connected scatterplot

Usually used to show how the relationship between 2 variables has changed over time.

Bubble

Like a scatterplot, but adds additional detail by sizing the circles according to a third variable.

XY heatmap

A good way of showing the patterns between 2 categories of data, less effective at showing fine differences in amounts.

Ranking

Use where an item's position in an ordered list is more important than its absolute or relative value. Don't be afraid to highlight the points of interest.

Example FT uses
Wealth, deprivation, league tables, constituency election results

Ordered bar

Standard bar charts display the ranks of values much more easily when sorted into order.

Ordered column

See above.

Ordered proportional symbol

Use when there are big variations between values and/or seeing fine differences between data is not so important.

Dot strip plot

Dots placed in order on a strip are a space-efficient method of laying out ranks across multiple categories.

Slope

Perfect for showing how ranks have changed over time or vary between categories.

Lollipop

Lollipops draw more attention to the data value than standard bar/column and can also show rank and value effectively.

Bump

Effective for showing changing rankings across multiple dates. For large datasets, consider grouping lines using colour.

Distribution

Show values in a dataset and how often they occur. The shape (or 'skew') of a distribution can be a memorable way of highlighting the lack of uniformity or equality in the data.

Example FT uses
Income distribution, population (ages/sex) distribution, revealing inequality

Histogram

The standard way to show a statistical distribution – keep the gaps between columns small to highlight the 'shape' of the data.

Dot plot

A simple way of showing the change or range (min/max) of data across multiple categories.

Dot strip plot

Good for showing individual values in a distribution, can be a problem when too many dots have the same value.

Barcode plot

Like dot strip plots, good for displaying all the data in a table, they work best when highlighting individual values.

Boxplot

Summarise multiple distributions by showing the median (centre) and range of the data

Violin plot

Similar to a box plot but more effective with complex distributions (data that cannot be summarised with simple average).

Population pyramid

A standard way for showing the age and sex breakdown of a population distribution; effectively, back to back histograms.

Cumulative curve

A good way of showing how unequal a distribution is: y axis is always cumulative frequency, x axis is always a measure.

Frequency polygons

For displaying multiple distributions of data. Like a regular line chart, best limited to a maximum of 3 or 4 datasets.

Beeswarm

Use to emphasise individual points in a distribution. Points can be sized to an additional variable. Best with medium-sized datasets

Change over Time

Give emphasis to changing trends. These can be short (intra-day) movements or extended series traversing decades or centuries. Choosing the correct time period is important to provide suitable context for the reader.

Example FT uses
Share price movements, economic time series, sectoral changes in a market

Line

The standard way to show a changing time series. If data are irregular, consider markers to represent data points.

Column

Columns work well for showing change over time – but usually best with only one series of data at a time.

Column + line timeline

A good way of showing the relationship over time between an amount (columns) and a rate (line).

Slope

Good for showing changing data as long as the data can be simplified into 2 or 3 points without missing a key part of story.

Area chart

Use with care – these are good at showing changes to total, but showing change in components can be very difficult.

Candlestick

Usually focused on day-to-day activity, these charts show opening/closing and high/low points of each day.

Fan chart (projections)

Use to show the uncertainty in future projections - usually this grows the further forward to projection.

Connected scatterplot

A good way of showing changing data for two variables whenever there is a relatively clear pattern of progression.

Calendar heatmap

A great way of showing temporal patterns (daily, weekly, monthly) – at the expense of showing precision in quantity.

Priestley timeline

Great when date and duration are key elements of the story in the data.

Circle timeline

Good for showing discrete values of varying size across multiple categories (eg earthquakes by continent).

Vertical timeline

Presents time on the Y axis. Good for displaying detailed time series that work especially well when scrolling on mobile.

Seismogram

Another alternative to the circle timeline for showing series where there are big variations in the data.

Streamgraph

A type of area chart; use when seeing changes in proportions over time is more important than individual values

Magnitude

Show size comparisons. These can be relative (just being able to see larger/bigger) or absolute (need to see fine differences). Usually these show a 'counted' number (for example, barrels, dollars or people) rather than a calculated rate or per cent.

Example FT uses
Commodity production, market capitalisation, volumes in general

Column

The standard way to compare the size of things. Must always start at 0 on the axis.

Bar

See above. Good when the data are not time series and labels have long category names.

Paired column

As per standard column but allows for multiple series. Can become tricky to read with more than 2 series.

Paired bar

See above.

Marimekko

A good way of showing the size and proportion of data across multiple categories.

Proportional symbol

Use when there are big variations between values and/or seeing fine differences between data is not so important.

Isotype (pictogram)

Excellent solution in some instances – use only with whole numbers (do not slice off an arm to represent a decimal).

Lollipop

Lollipop charts draw more attention to the data value than standard bar/column – does not have to start at zero (but preferable).

Radar

A space-efficient way of showing value of multiple variables– but make sure they are organised in a way that makes sense to reader.

Parallel coordinates

An alternative to radar charts – again, the arrangement of the variables is important. Usually benefits from highlighting values.

Bullet

Good for showing a measurement against the context of a target or performance range.

Grouped symbol

An alternative to bar/column charts when being able to count data or highlight individual elements is useful.

Part-to-whole

Show how a single entity can be broken down into its component elements. If the reader's interest is solely in the size of the components, consider a magnitude-type chart instead.

Example FT uses
Fiscal budgets, company structures, national election results

Stacked column/bar

A simple way of showing part-to-whole relationships but can be difficult to read with more than a few components.

Marimekko

A good way of showing the size and proportion of data at the same time – as long as the data are not too complicated.

Pie

A common way of showing part-to-whole data – but be aware that it's difficult to accurately compare the size of the segments.

Donut

Similar to a pie chart – but the centre can be a good way of making space to include more information about the data (eg total).

Treemap

Use for hierarchical part-to-whole relationships; can be difficult to read when there are many small segments.

Voronoi

A way of turning points into areas – any point within each area is closer to the central point than any other centroid.

Arc

A hemicycle, often used for visualising parliamentary composition by number of seats.

Gridplot

Good for showing % information, they work best when used on whole numbers and work well in small multiple layout form.

Venn

Generally only used for schematic representation.

Waterfall

Can be useful for showing part-to-whole relationships where some of the components are negative.

Spatial

Aside from locator maps only used when precise locations or geographical patterns in data are more important to the reader than anything else.

Example FT uses
Population density, natural resource locations, natural disaster risk/impact, catchment areas, variation in election results

Basic choropleth (rate/ratio)

The standard approach for putting data on a map – should always be rates rather than totals and use a sensible base geography.

Proportional symbol (count/magnitude)

Use for totals rather than rates – be wary that small differences in data will be hard to see.

Flow map

For showing unambiguous movement across a map.

Contour map

For showing areas of equal value on a map. Can use deviation colour schemes for showing +/- values

Equalised cartogram

Converting each unit on a map to a regular and equally-sized shape – good for representing voting regions with equal value.

Scaled cartogram (value)

Stretching and shrinking a map so that each area is sized according to a particular value.

Dot density

Used to show the location of individual events/locations – make sure to annotate any patterns the reader should see.

Heat map

Grid-based data values mapped with an intensity colour scale. As choropleth map – but not snapped to an admin/political unit.

Flow

Show the reader volumes or intensity of movement between two or more states or conditions. These might be logical sequences or geographical locations.

Example FT uses
Movement of funds, trade, migrants, lawsuits, information; relationship graphs.

Sankey

Shows changes in flows from one condition to at least one other; good for tracing the eventual outcome of a complex process.

Waterfall

Designed to show the sequencing of data through a flow process, typically budgets. Can include +/- components.

Chord

A complex but powerful diagram which can illustrate 2-way flows (and net winner) in a matrix.

Network

Used for showing the strength and inter-connectedness of relationships of varying types.

Ranking

Use where an item's position in an ordered list is more important than its absolute or relative value. Don't be afraid to highlight the points of interest.

Example FT uses
Wealth, deprivation, league tables, constituency election results

Ordered bar

Standard bar charts display the ranks of values much more easily when sorted into order.

Ordered column

See above.

Ordered proportional symbol

Use when there are big variations between values and/or seeing fine differences between data is not so important.

Dot strip plot

Dots placed in order on a strip are a space-efficient method of laying out ranks across multiple categories.

Slope

Perfect for showing how ranks have changed over time or vary between categories.

Lollipop

Lollipops draw more attention to the data value than standard bar/column and can also show rank and value effectively.

Distribution

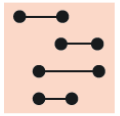
Show values in a dataset and how often they occur. The shape (or 'skew') of a distribution can be a memorable way of highlighting the lack of uniformity or equality in the data.

Example FT uses
Income distribution, population (age/sex) distribution, revealing inequality

Histogram

The standard way to show a statistical distribution - keep the gaps between columns small to highlight the 'shape' of the data.

Dot plot



A simple way of showing the change or range (min/max) of data across multiple categories.

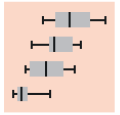
Dot strip plot

Good for showing individual values in a distribution, can be a problem when too many dots have the same value.

Barcode plot

Like dot strip plots, good for displaying all the data in a table, they work best when highlighting individual values.

Boxplot



Summarise multiple distributions by showing the median (centre) and range of the data

Violin plot

Similar to a box plot but more effective with complex distributions (data that cannot be summarised with simple average).

Change over Time

Give emphasis to changing trends. These can be short (intra-day) movements or extended series traversing decades or centuries: Choosing the correct time period is important to provide suitable context for the reader.

Example FT uses
Share price movements, economic time series, sectoral changes in a market

Line

The standard way to show a changing time series. If data are irregular, consider markers to represent data points.

Column

Columns work well for showing change over time - but usually best with only one series of data at a time.

Column + line timeline

A good way of showing the relationship over time between an amount (columns) and a rate (line).

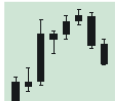
Slope

Good for showing changing data as long as the data can be simplified into 2 or 3 points without missing a key part of story.

Area chart

Use with care – these are good at showing changes to total, but seeing change in components can be very difficult.

Candlestick



Usually focused on day-to-day activity, these charts show opening/closing and high/low points of each day.

Magnitude

Show size comparisons. These can be relative (just being able to see larger/bigger) or absolute (need to see fine differences). Usually these show a 'counted' number (for example, barrels, dollars or people) rather than a calculated rate or per cent.

Example FT uses
Commodity production, market capitalisation, volumes in general

Column

The standard way to compare the size of things. Must always start at 0 on the axis.

Bar

See above. Good when the data are not time series and labels have long category names.

Paired column

As per standard column but allows for multiple series. Can become tricky to read with more than 2 series.

Paired bar

See above.

Marimekko

A good way of showing the size and proportion of data at the same time – as long as the data are not too complicated.

Proportional symbol

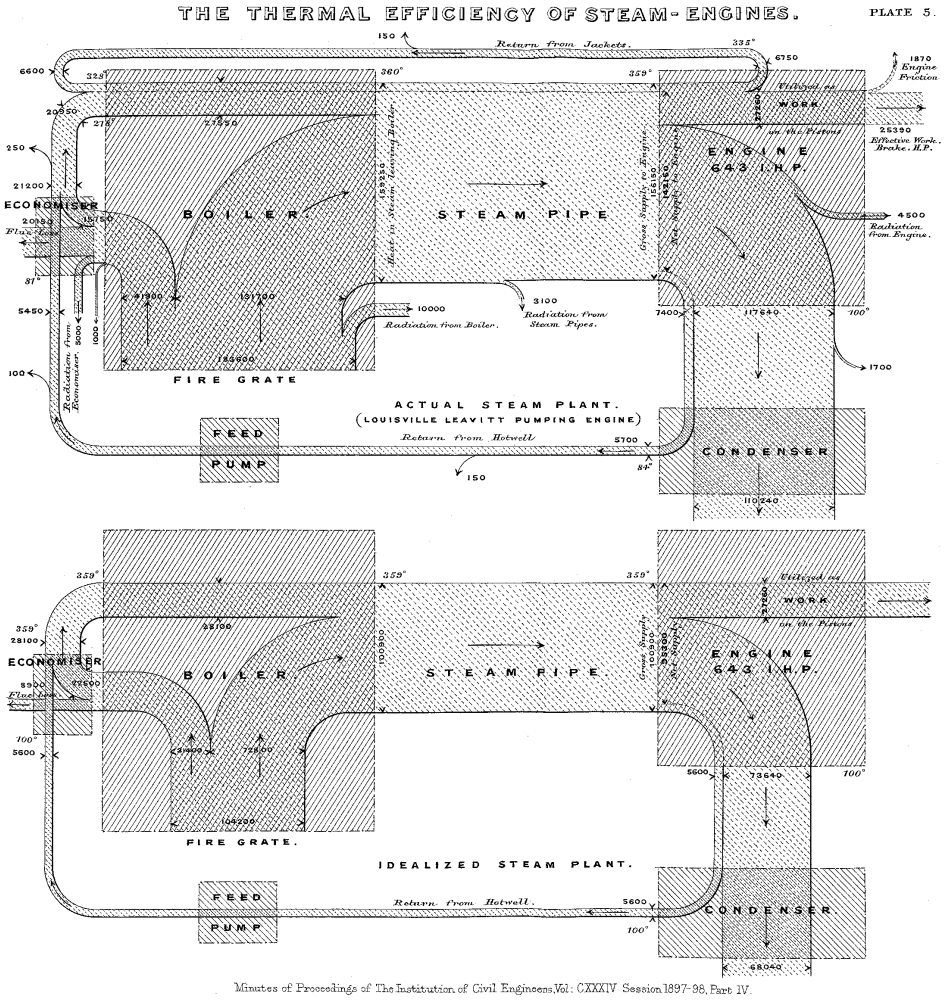
Use when there are big variations between values and/or seeing fine differences between data is not so

Specialist Charts- Sankey

Simple charts most efficient, but not visually very exciting. Consider experimenting with more advanced charts, especially where looking for more complex data analysis.

Sankey diagrams are a type of flow map where width of flow is proportional to volume. Invented by Irish engineer Sankey in late 19th century (note French cartographer Minard made similar innovations a century before).

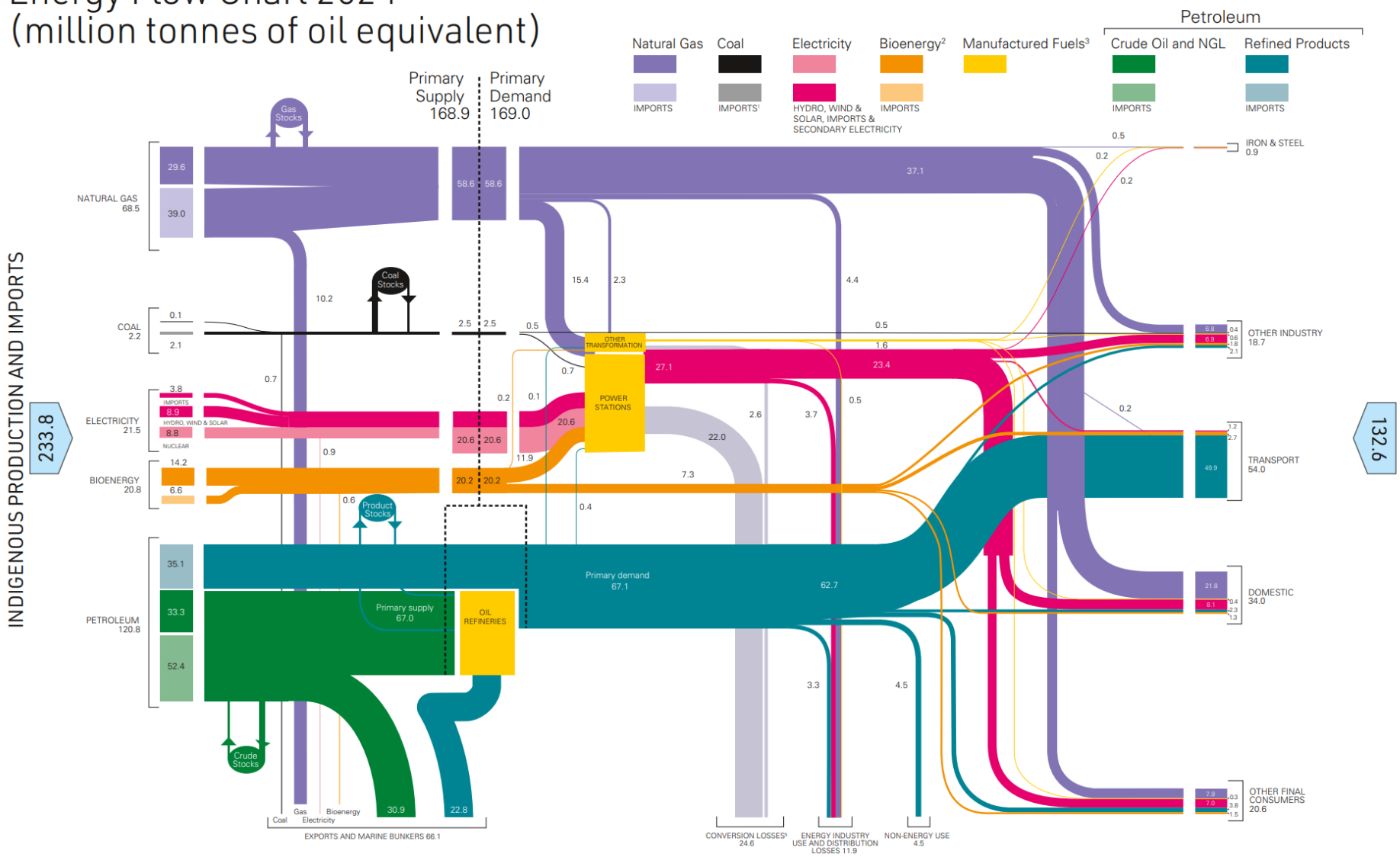
Very useful for understanding energy and efficiency flows. Relatively complex for general audience.



Original Sankey diagram on efficiency of steam engine 1898.

UK Energy Production and Consumption Sankey: https://www.gov.uk/government/collections/energy-flow-charts

Energy Flow Chart 2024
(million tonnes of oil equivalent)



FOOTNOTES:
1. Coal imports, exports and power stations include manufactured fuels.
2. Bioenergy is renewable energy made from material of recent biological origin derived from plant or animal matter.
3. Includes heat sold.
4. Includes non-energy use.
5. Conversion losses from power stations apply to thermal sources (coal, oil, gas, nuclear and bioenergy) only; there are no such losses from non-thermal sources (hydro, wind and solar).

This flowchart has been produced using the style of balance and figures in the 2025 Digest of UK Energy Statistics, Table 1.1. (gross calorific values basis)

Department for Energy Security & Net Zero

OFFICIAL STATISTICS Accredited

Specialist Charts- Calendar Heatmap

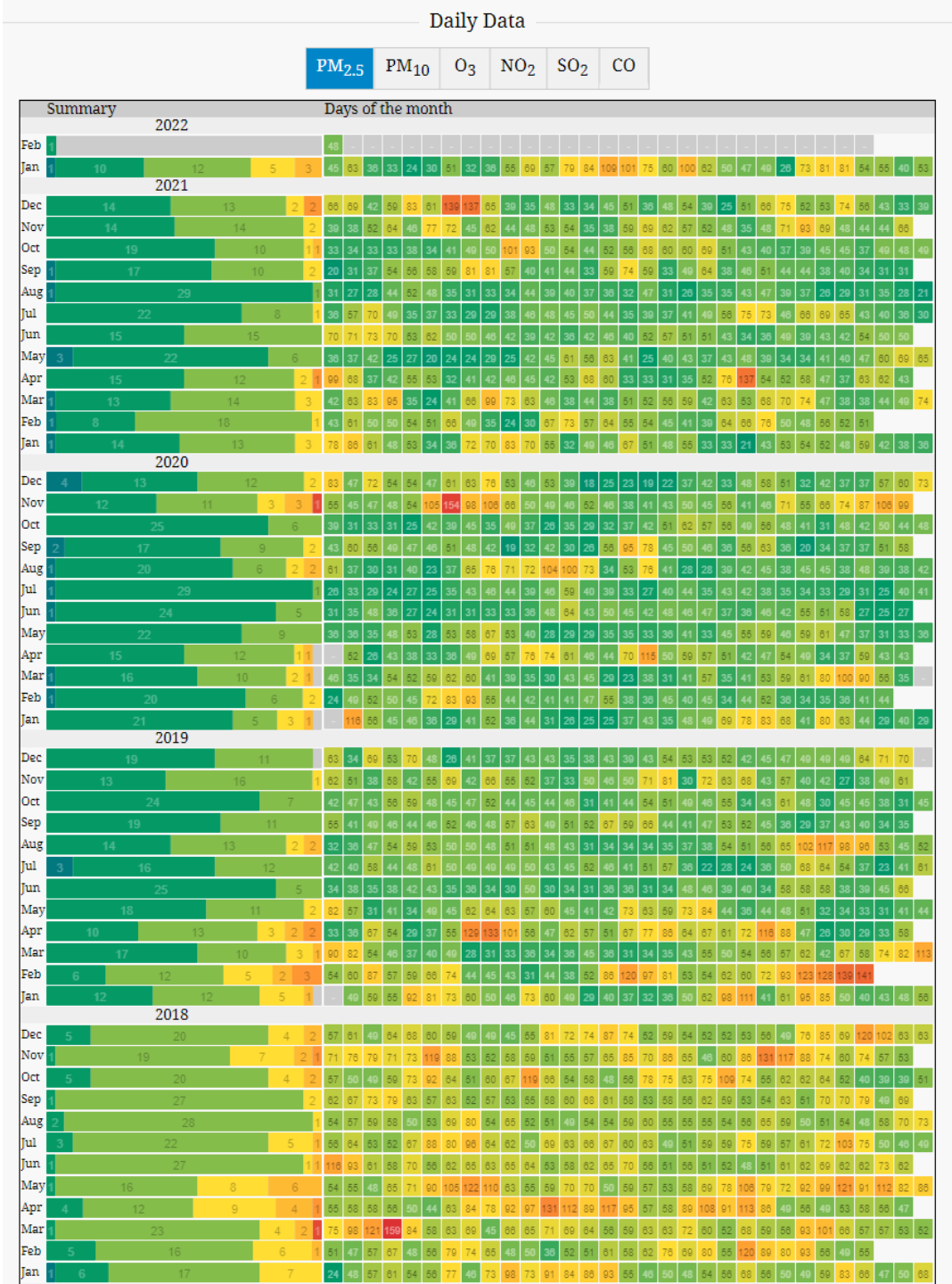
Similar to correlation matrix plots used in statistical software (e.g. R). Simple and effective way to summarise time-series data where daily averages are a key measure.

Good example is air quality data, which is dependent on highly variable weather patterns as well as local emission sources (transport, energy)-

London PM2.5 Daily Average

Chart from world air quality index- http://aqicn.org/

AQI	Air Pollution Level
0 - 50	Good
51 -100	Moderate
101-150	Unhealthy for Sensitive Groups
151-200	Unhealthy
201-300	Very Unhealthy
300+	Hazardous



Delhi PM2.5 Daily Average

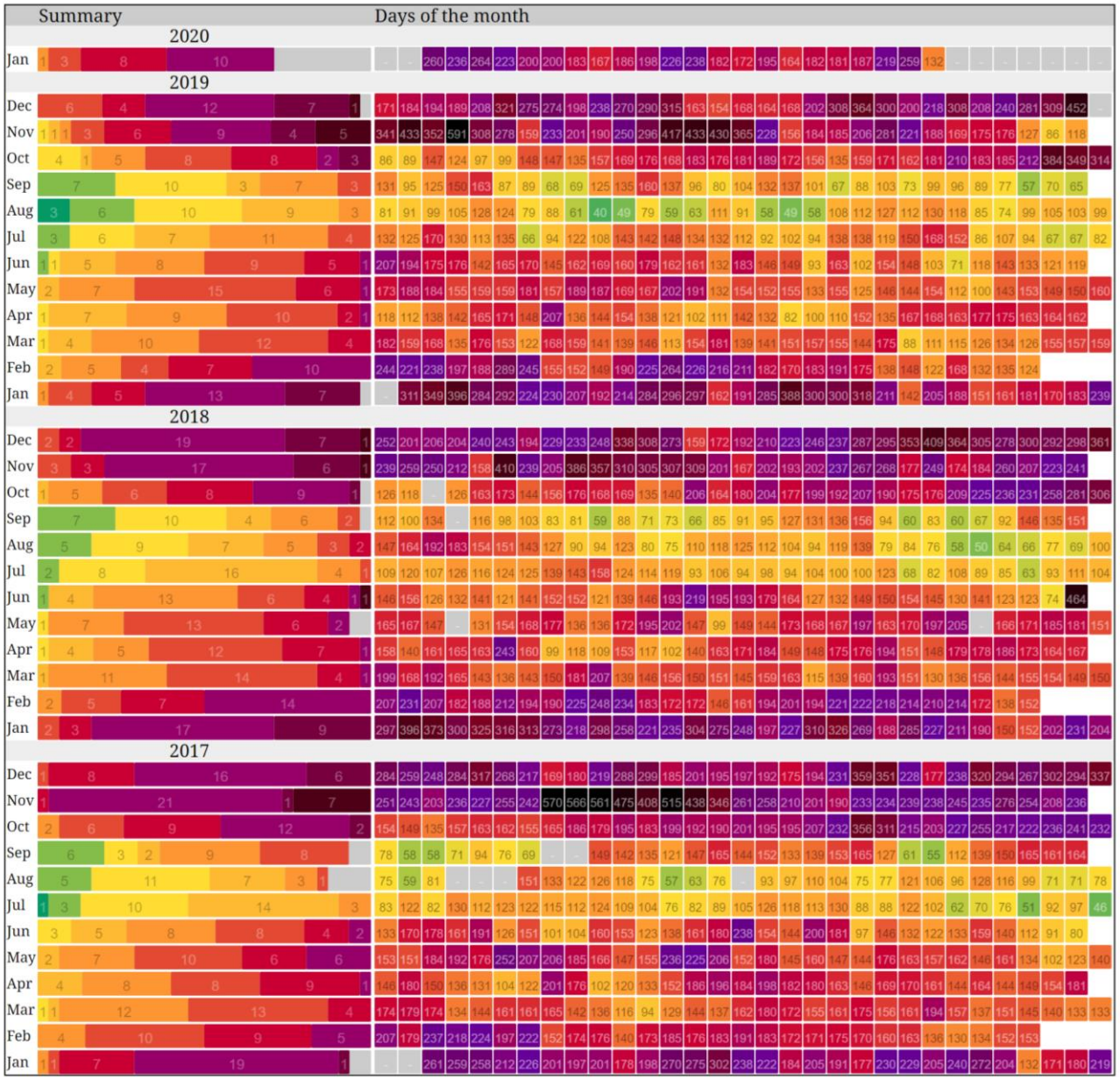


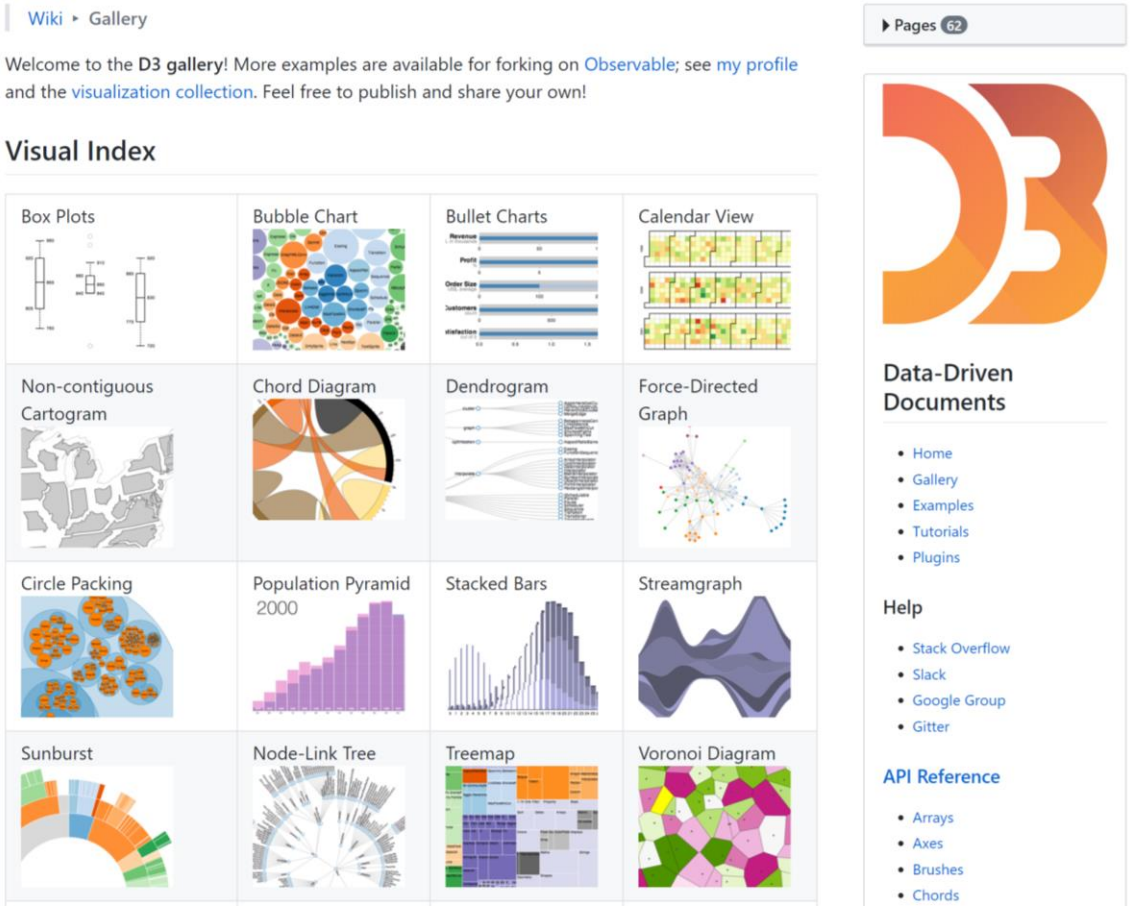
Chart from world air quality index- http://aqicn.org/

(Note Delhi data based on single sensor, rather than city average)

AQI	Air Pollution Level
0 - 50	Good
51 -100	Moderate
101-150	Unhealthy for Sensitive Groups
151-200	Unhealthy
201-300	Very Unhealthy
300+	Hazardous

Templates for Charts

Generally easy to achieve variety of chart types using JavaScript, based on existing templates. The library we will be using in the practical is Chart.js. D3 is another popular tool-



Overview

1. Intro
2. Choosing Chart Types
3. **Charts, Maps & Animation**

Practical-

5. Creating Charts in JavaScript

Charts and Maps Complementary

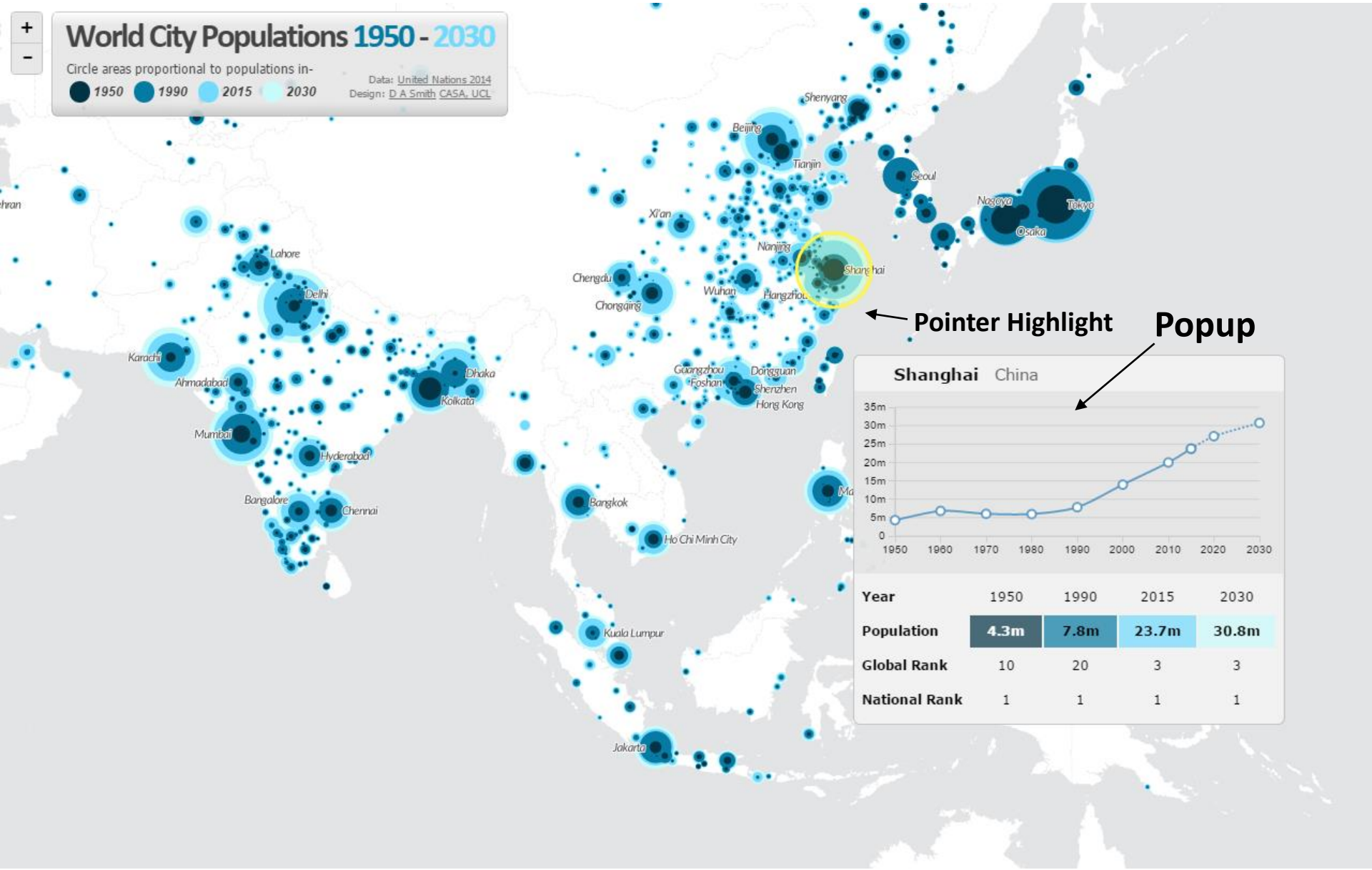
Given the different strengths of maps and charts, using them in combination can be effective.

Typical examples-

Adding timeline charts to maps to include temporal dimension

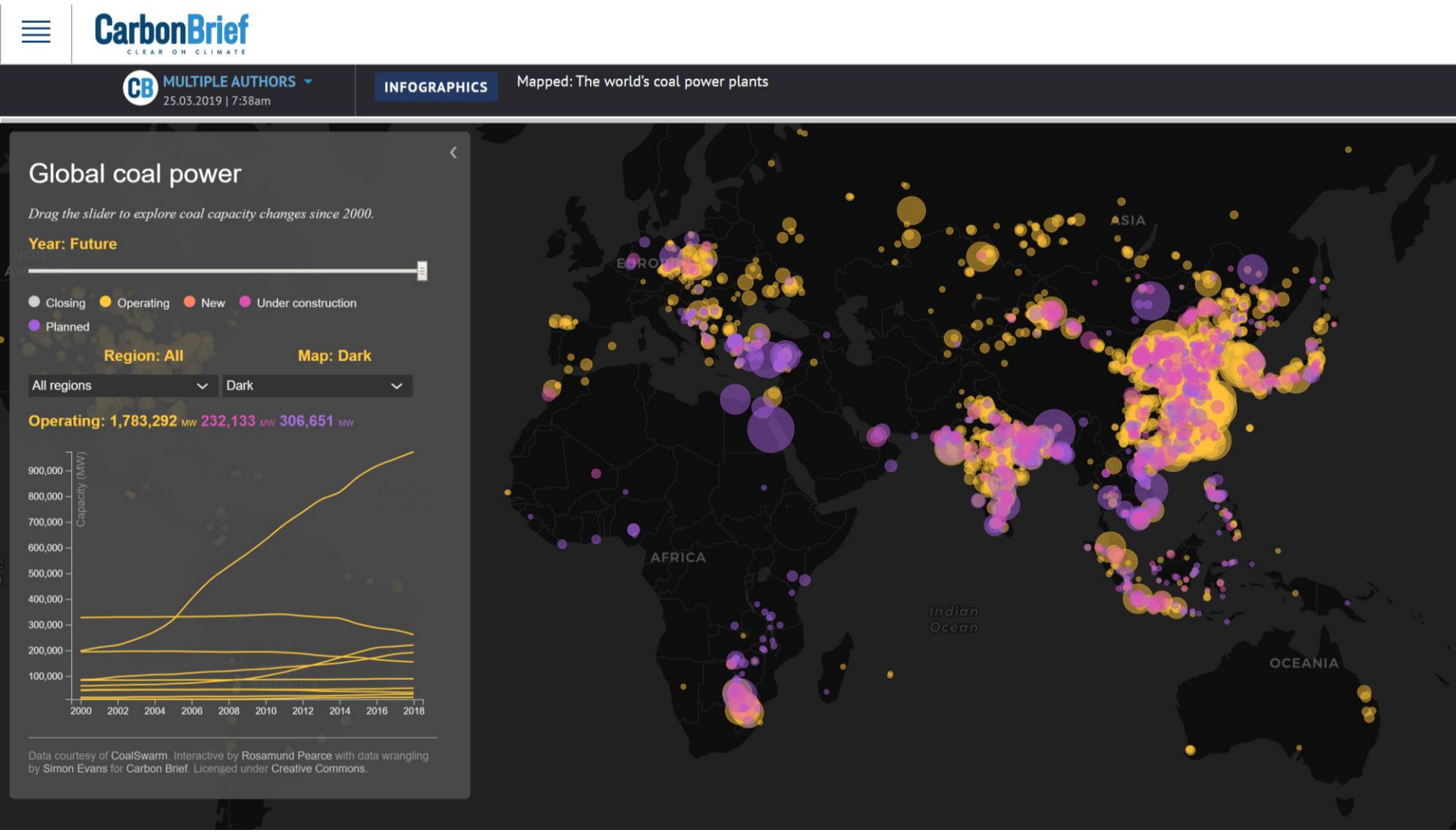
‘Brushing’ interactivity- where chart selections highlight and change map navigation.

Popup Chart Example- Timeseries and Figures



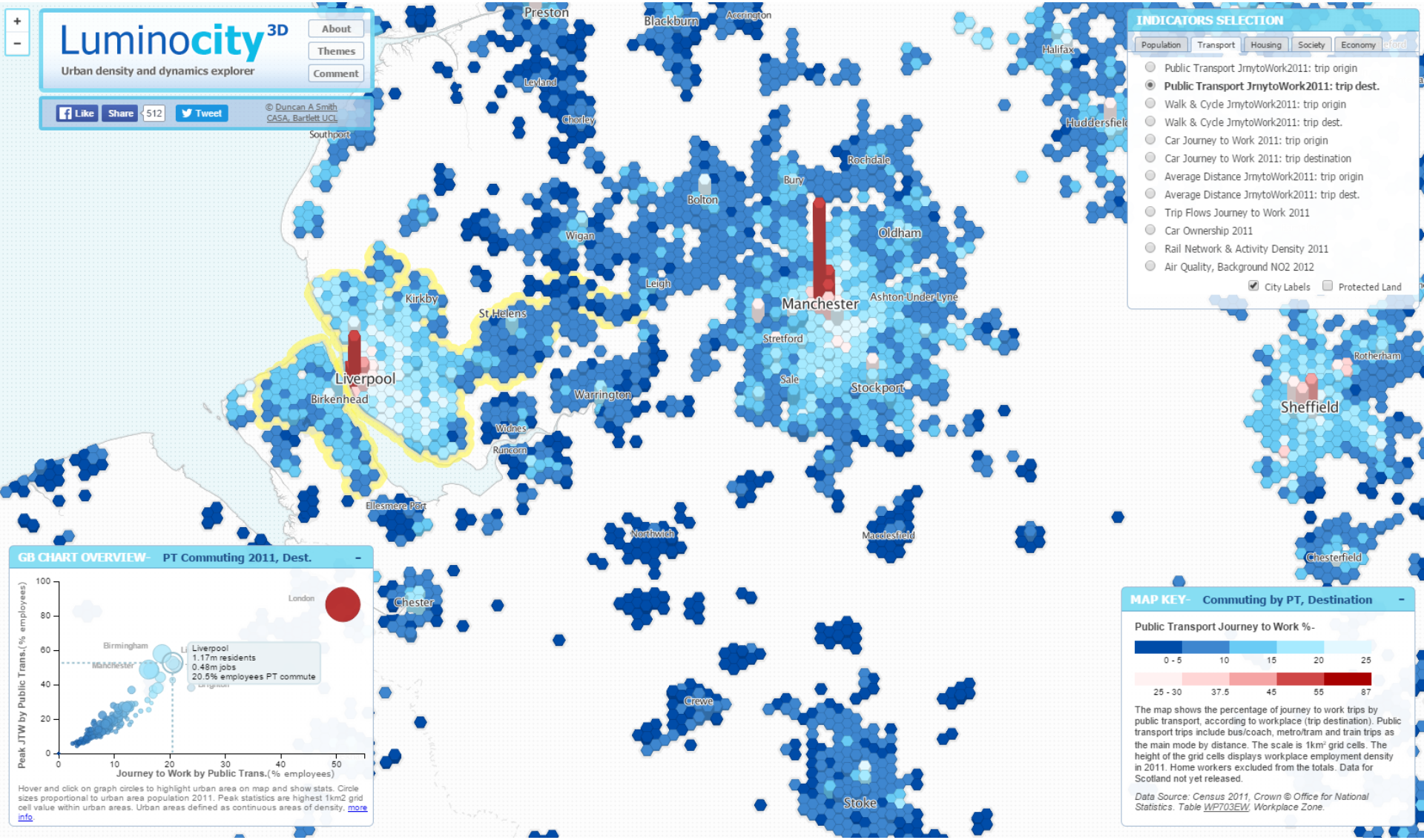
https://luminocity3d.org/WorldCity/

Linechart Timeline Example with Regional Totals



https://www.carbonbrief.org/mapped-worlds-coal-power-plants

Brushing Example- Scatterplot and Map Link



https://luminocity3d.org/Transport.html

Animated Charts

Transitions are smooth changes from one state to another. Allow greater visual continuity, help viewers understand changes.

We discussed animated transitions in the context of online maps last week. Also relevant for charts.

Common Animated Transitions for Online Interactive Charts

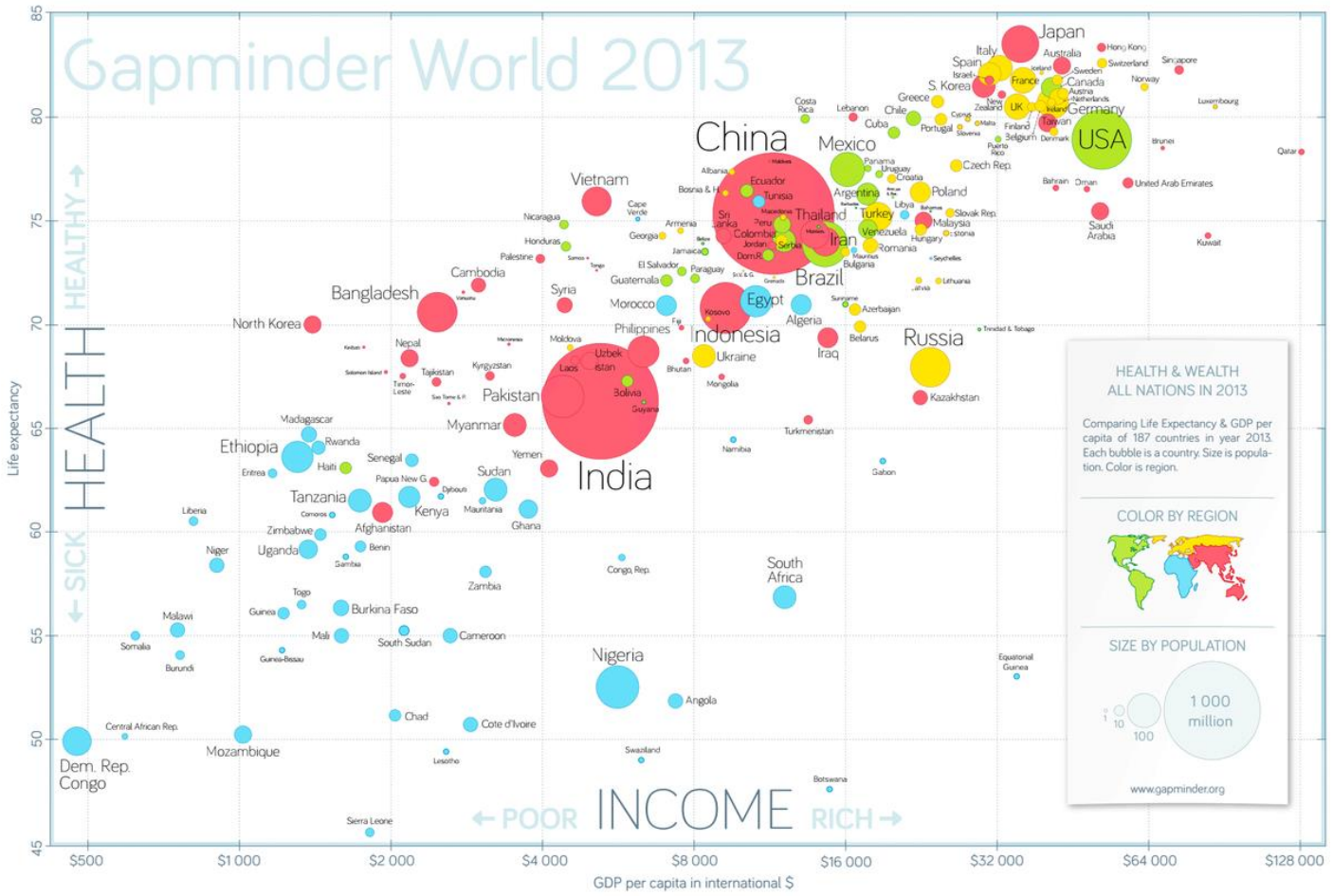
Time series animations;

Filtering data/changing the data variable;

Changing chart types.

Time Series Animation

Time is animated, usually through a play button. Gapminder (Hans Rosling) is a famous example of demographic change over time in a bubble chart (https://www.gapminder.org/tools/)-

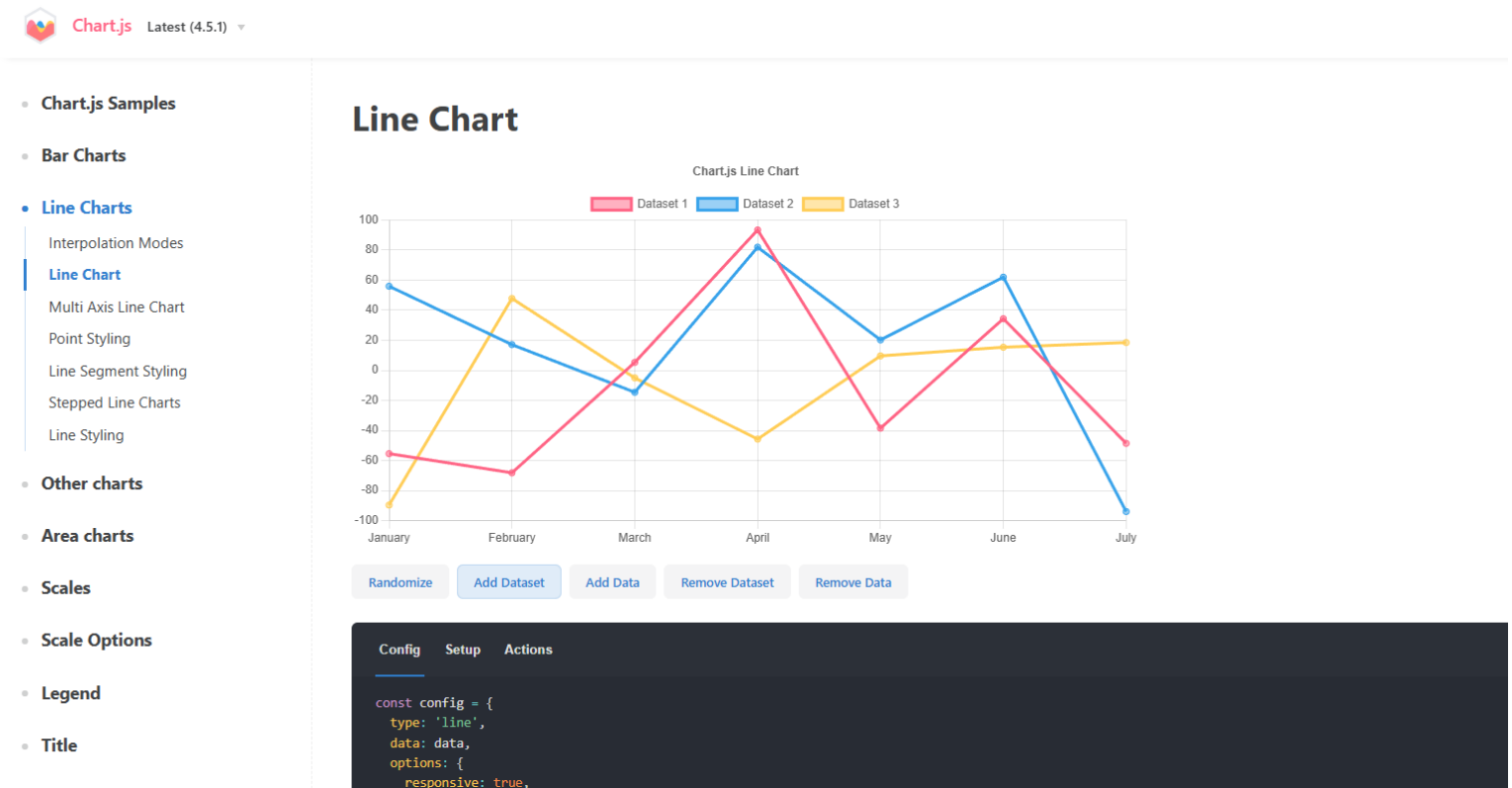


INCOME DATA: World Bank's GDP per capita, PPP (constant 2011 international $), with a few additions by Gapminder. The x-axis uses a log-scale so that doubling incomes show the same distance on all levels. LIFE EXPECTANCY & POPULATION DATA: UN Population Division. INTERACTIVE TOOL: a free interactive version of this chart is available online at www.gapminder.org/world, which lets you play historic time series & compare other indicators. LICENSE: Our charts are freely available under Creative Commons Attribution License. Please copy, share, modify, integrate and even sell them, as long as you mention "Based on a free chart from www.gapminder.org".

Data Change Animations

Fairly standard in online maps to be able to smoothly filter data or change the variables being charted. Examples from Chart.js library-
https://www.chartjs.org/docs/latest/samples/bar/horizontal.html

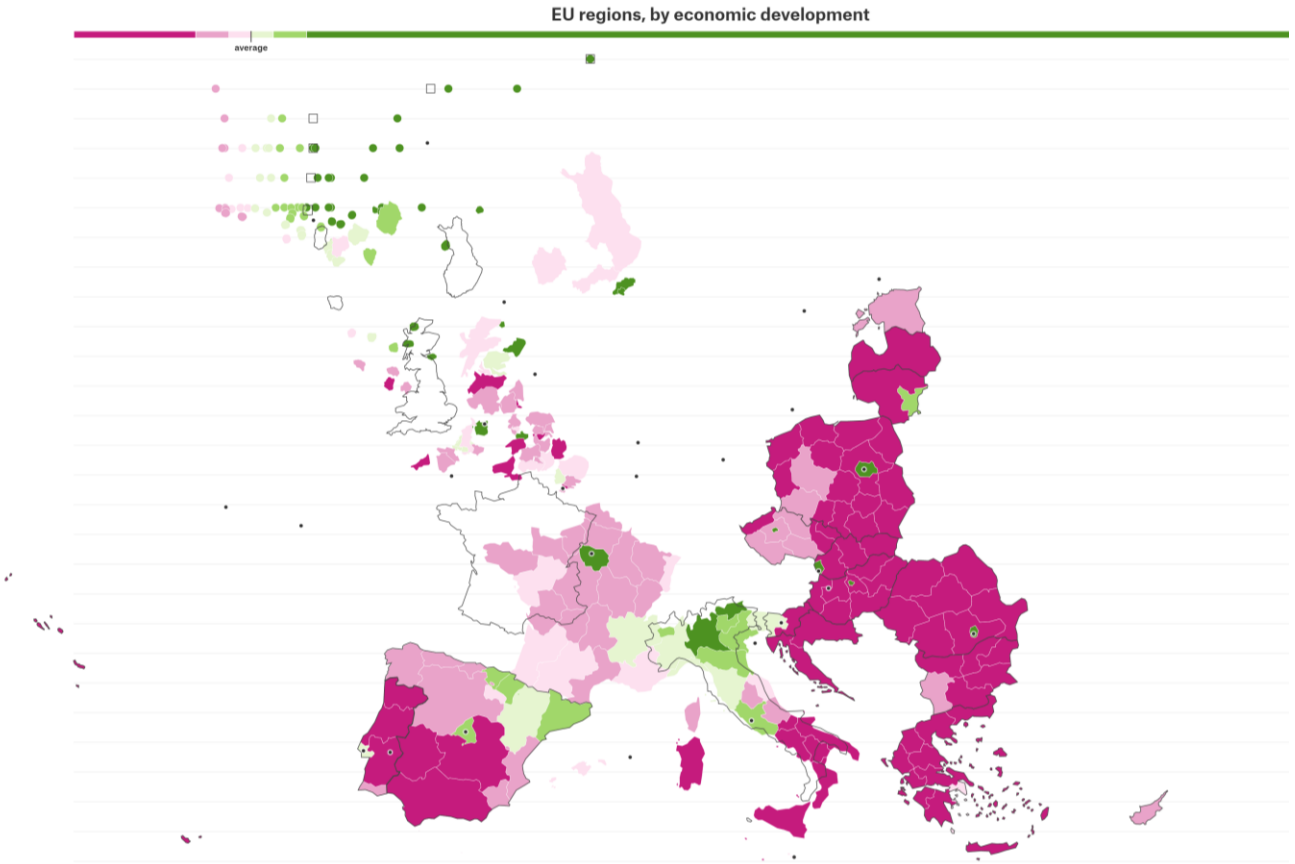
https://www.chartjs.org/docs/latest/samples/line/line.html



Changing Chart Type Animation

Arguably more flashy than useful transition, but can be visually very impressive. Example of map to chart animation using D3-

https://pudding.cool/2019/04/eu-regions/



Conclusions

Charts Best Way to Show Magnitude and Change over Time
Maps best for spatial patterns but have weaknesses. Situations where charts will be a better choice, or can be used in combination with maps.

Simple Charts more Cognitively Efficient
Length and position comparisons on a single scale easiest and most accurate estimates from users.

Lots of Innovation in more Specialist Charts, Interactive
Lots of sophisticated charts possible where more in-depth variable exploration desired. Experiment with these, but recognise more challenging for some viewers.

Good Online JavaScript Tools for Chart Creation
Lots of libraries available, from cloud service tools, to sophisticated libraries such as D3 that give you total control.

Group Discussion

There are many different types of charts and statistical visualisations, leading to important design decisions regarding when to use charts and/or maps; and which types of charts to use depending on the data and research questions.

The Financial Times have developed an impressive reputation for high quality data visualisation and data journalism (see for example their excellent charts on tracking Covid-19 across the globe, or excellent work by John Burn-Murdoch (@jburnmurdoch). The Financial Times has developed a Visual Vocabulary describing different chart types and when to use them- https://github.com/Financial-Times/chart-doctor/blob/main/visual-vocabulary/Visual-vocabulary-en.pdf

What are the main criteria developed here for choosing between different chart types? Are any of these chart types new to you?

2. Recent visualisation tools (such as D3) make it easier to develop more sophisticated data visualisation types, such as Sankey charts- https://www.gov.uk/government/collections/energy-flow-charts
Chord diagrams- https://observablehq.com/@d3/chord-diagram
And Ridgeline Charts amongst others- https://www.d3-graph-gallery.com/ridgeline.html

What are the advantages of developing these rich and spectacular statistical visualisations? What might some of the challenges be with more sophisticated statistical visualisations?

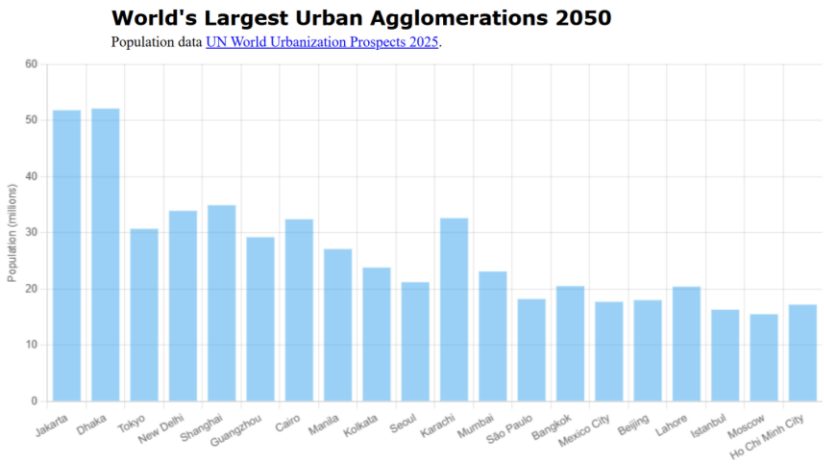
Practical

Going to learn how to create charts in JavaScript.

There are lots of good libraries available (e.g. HighCharts, Plotly, Echarts…). We are going to use a popular library called Chart.js, which can create basic charts very easily.

For more advanced chart types, libraries such as D3, and cloud service tools such as Flourish and Tableau are popular.

Chart.js



Homework: Choose Dataset and Visualisation Type for Individual Visualisation

Look for a suitable dataset you would like to use for your Individual Visualisation. Can fit with the 'Urban Analytics' theme. Good places to look: data.gov.uk; city datastores; EU data; OpenStreetMap; Digimap…

What would be the best visualisation type for your dataset? A chart? A map? What kind of map/chart?

What research questions can be explored with this dataset? Could interactive functionality help with these research questions?

We want everyone to post on Slack with their Individual Visualisation idea. Should include – dataset name; visualisation type; library/tool you are planning to use.