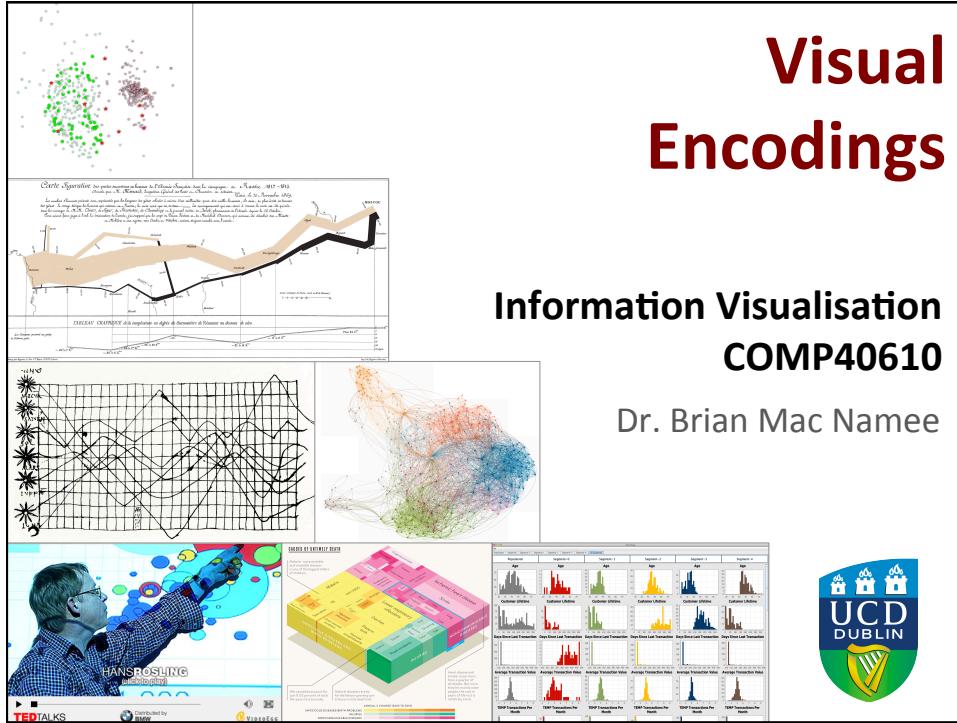


Visual Encodings

Information Visualisation
COMP40610

Dr. Brian Mac Namee



Origins

This course curates material from multiple
online and published sources

When this is the case full citations will be given

Agenda

In this lecture we will cover

- Jacques Bertin and the Semiology of Graphics
- Visual encodings & data dimensions
- Using visual encodings effectively

JACQUES BERTIN &
THE SEMIOLOGY OF GRAPHICS

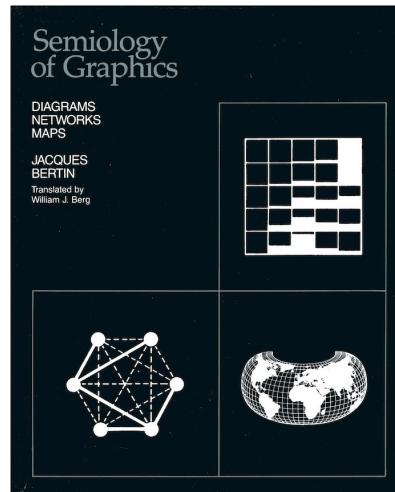


Jacques Bertin

Semiology of Graphics

The **Semiology of Graphics: diagrams networks maps** was first published in French in 1967

- First proposal of a semiology (or semiotics) of information visualisation
- Translated into English for the first time in 1980s, rereleased in 2014



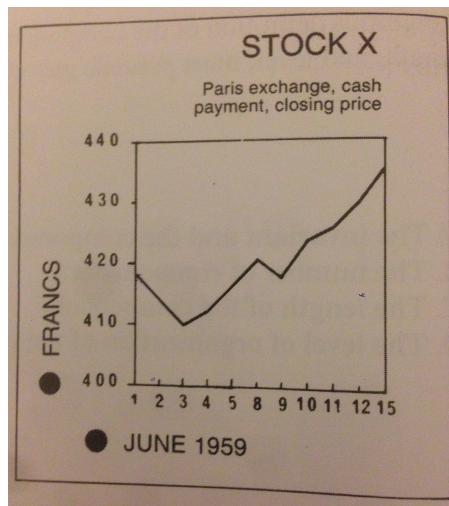
符号学 invariant 不变

Semiology of Graphics

Bertin describes the key characteristics of an information visualisation as a display of the relationship between an invariant and its components:

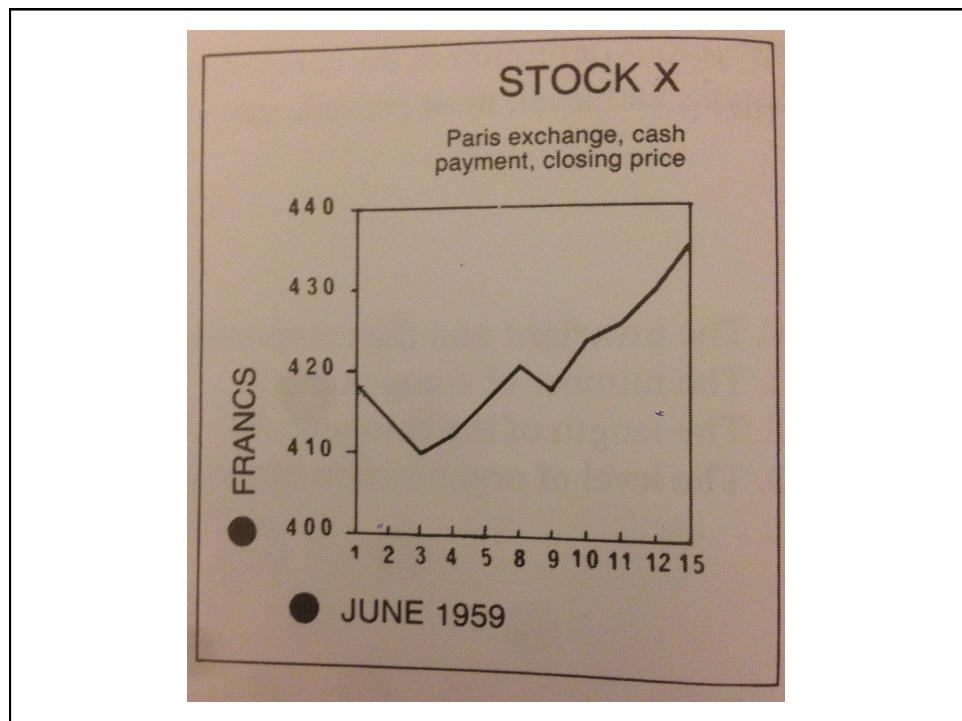
- The **invariant** is the complete and invariable notion common to all the data
- The **components** are the variational concepts

Semiology of Graphics

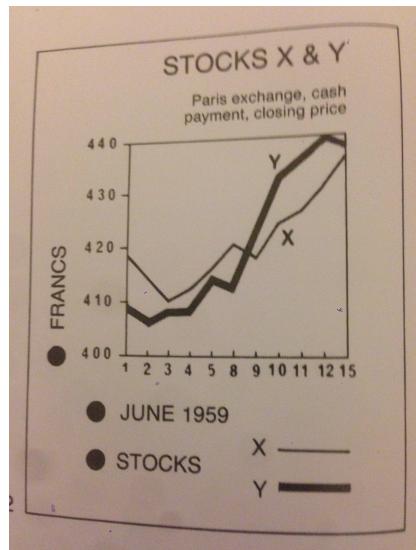


The trend of stock X on the Paris exchange

- **Invariant:** quotation in francs for stock X, cash payment, closing price on the Paris exchange
- **Components:** quantities (in francs), time (in days)

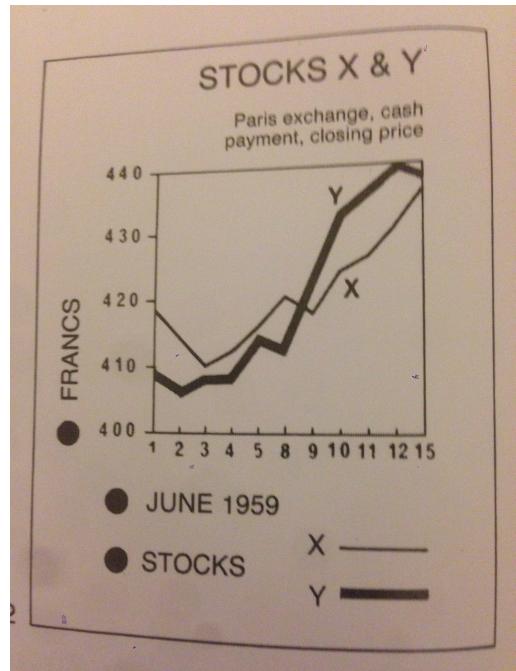


Semiology of Graphics

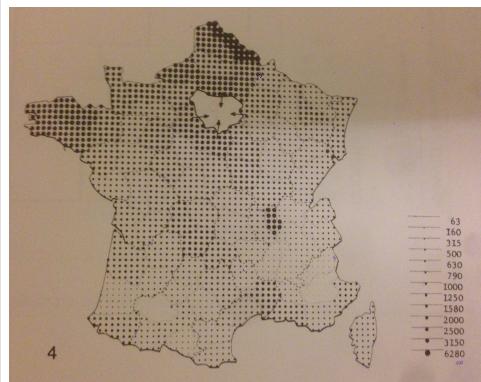


Comparison of the trends of stocks X and Y on the Paris exchange

- **Invariant:** quotation in francs, cash payment, closing price on the Paris exchange
- **Components:** quantities (in francs), time (in days), different stocks (X and Y)

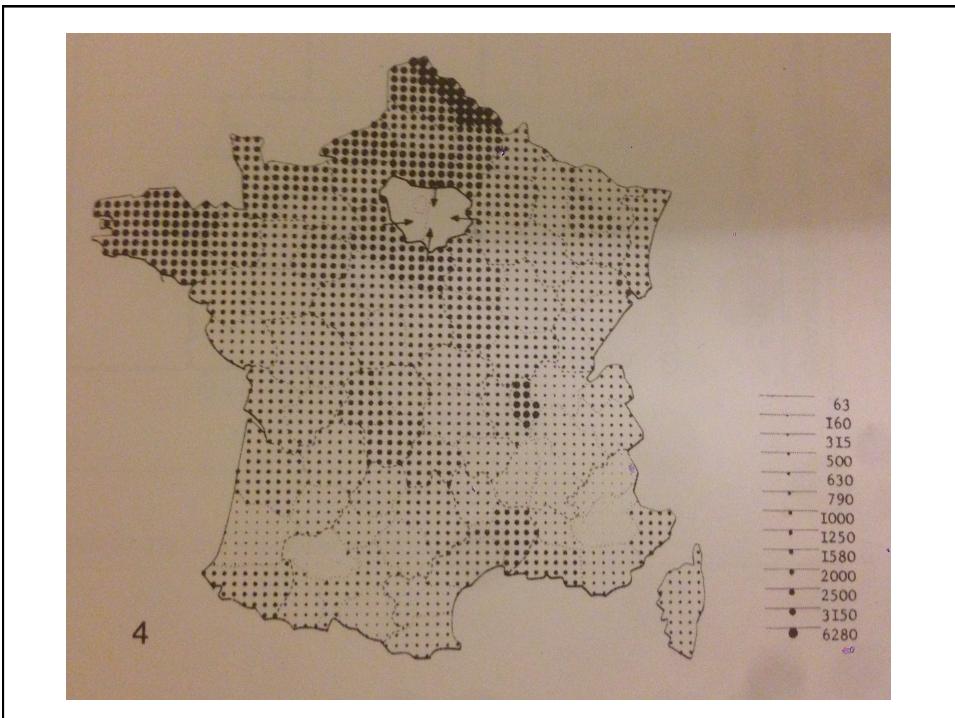


Semiology of Graphics



Distribution per 100 persons born outside the Paris area but residing there in 1962, according to their department of birth

- **Invariant:** a person living in Paris, born in the provinces
- **Components:** population, department

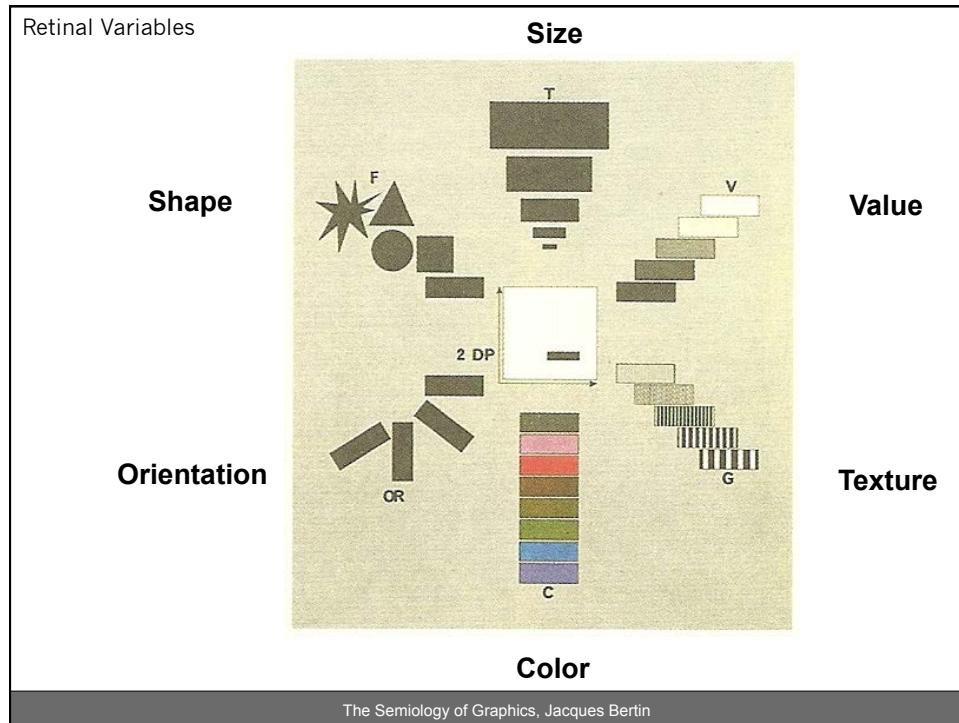


retinal 视网膜

Semiology of Graphics

Bertin defined a small set of encodings that are possible in information visualisations: 2 dimensions of the plane, and 8 **retinal variables**:

- Size
- (Colour) Value
- Texture
- Colour (Hue)
- Orientation
- Shape



Semiology of Graphics

Bertin also distinguishes three ways to represent things in the plane

- Points
- Lines
- Areas

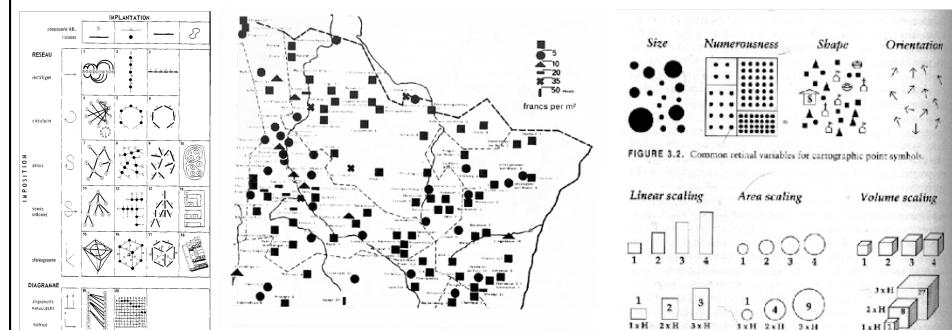
Also goes on to make recommendations about what variables work best in different scenarios

	Points	Lines	Areas	Best to show
Shape		possible, but too weird to show	cartogram	qualitative differences
Size			cartogram	quantitative differences
Color Hue				qualitative differences
Color Value				quantitative differences
Color Intensity				qualitative differences
Texture				qualitative & quantitative differences

The Semiology of Graphics, Jacques Bertin

Semiology of Graphics

There is lots more in the Semiology of Graphics (particularly around networks and maps) and it is well worth reading



VISUAL ENCODINGS & DATA DIMENSIONS

Basic Types Of Data

Nominal (qualitative): no inherent order

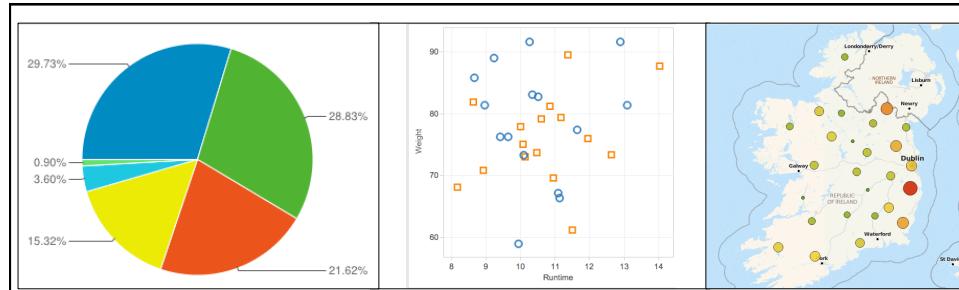
- city names, types of diseases, ...

Ordinal (qualitative): ordered, but not at measurable intervals

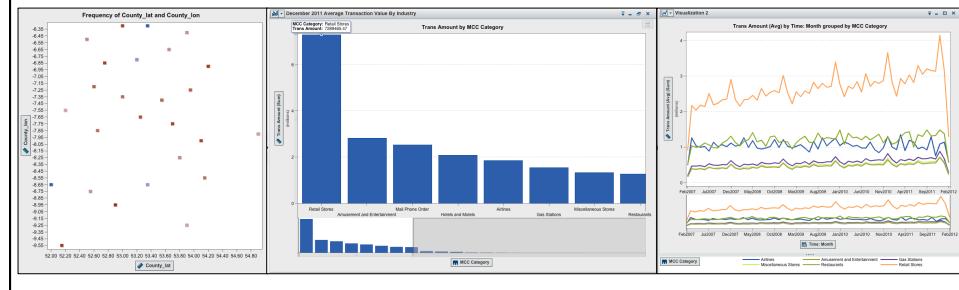
- first, second, third, ...
- cold, warm, hot
- Mon, Tue, Wed, Thu ...

Interval (quantitative): integers or reals

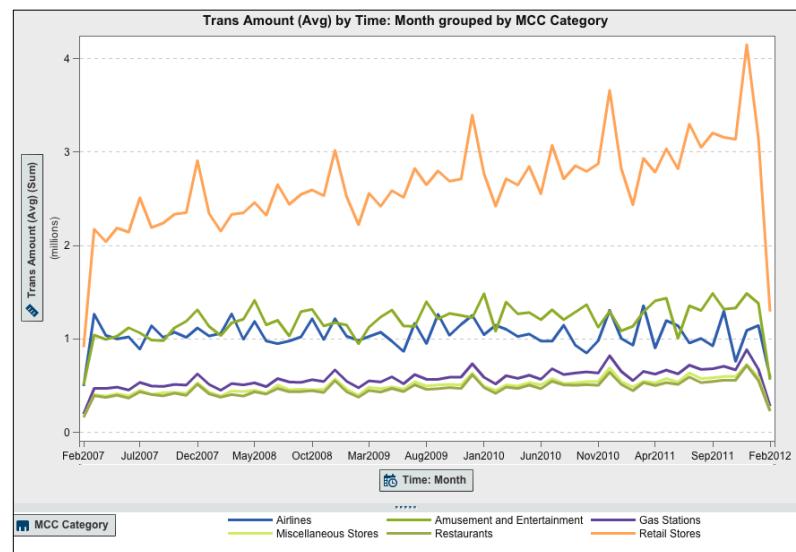
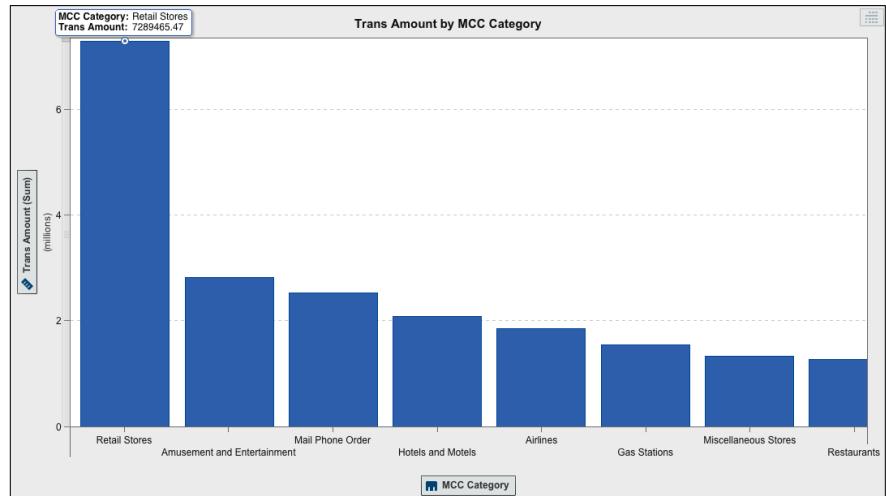
- 1, 2, 7
- 3.4, 78.6, - 12.4

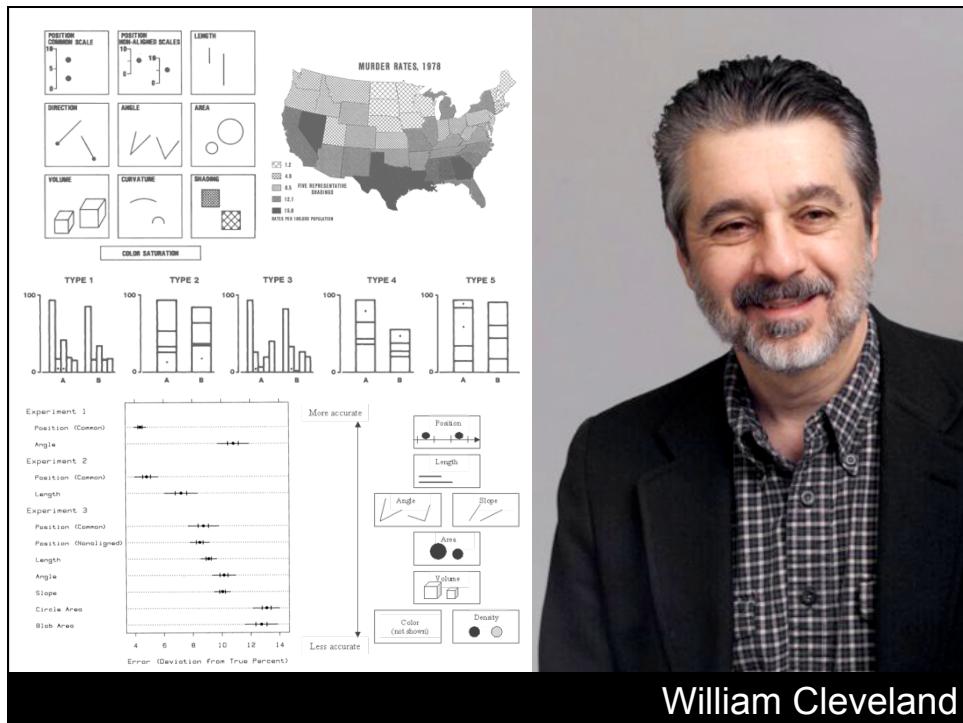
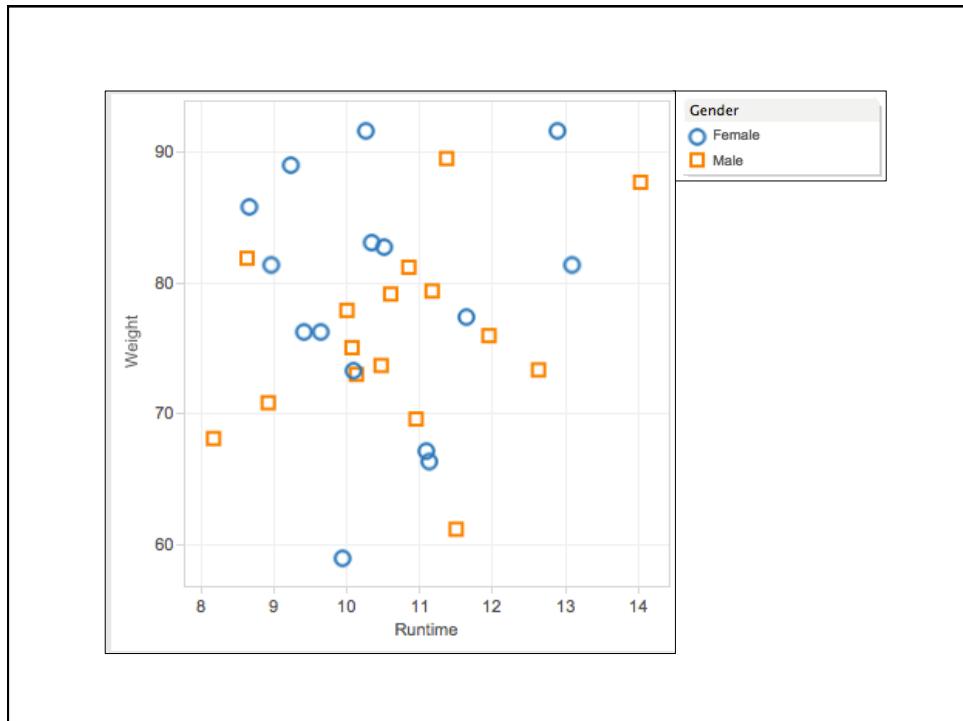


Fundamentally data visualisation is about mapping data dimensions to visual encodings



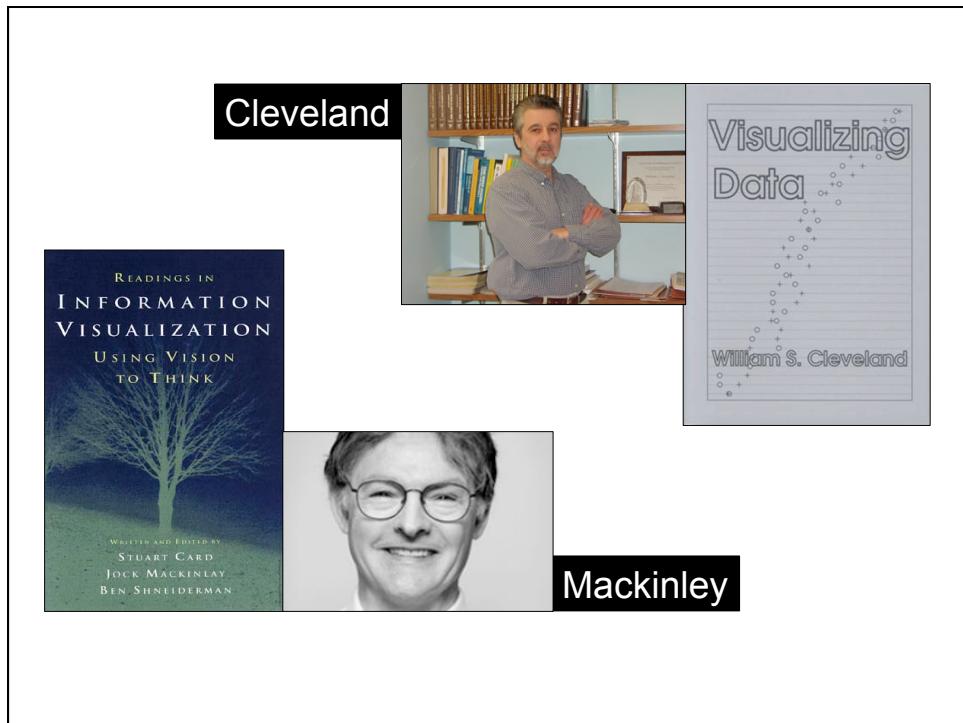
Encoding	Example	Ordered	Values
Position	● ● ●	Yes	Many
Length	█ █ █	Yes	Many
Angle	↙ ↘ ↗ ↖	Yes	Moderate
Area	• • ● ● ●	Yes	Moderate
Symbol	● ■ ★	No	Few
Colour Hue	■ ■ ■	No	Few
Colour Value	■ ■ ■ ■	Yes	Few
Pattern	— — — —	No	Few



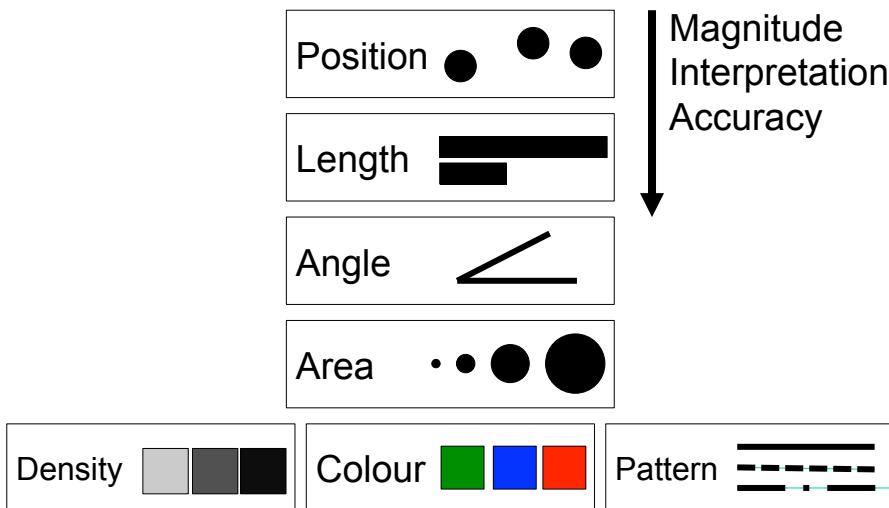




Jock MacKinlay



Ease of Perception Of Encodings



Graphical perception: Theory, experimentation, and application to the development of graphical methods, W S Cleveland, R McGill,
Journal of the American Statistical Association 01/1984; 79:531-554.

https://www.researchgate.net/publication/229099907_Graphical_perception_Theory_experimentation_and_application_to_the_development_of_graphical_methods

USING VISUAL ENCODINGS EFFECTIVELY

Using Visual Encodings Effectively

These are some things you should think about to help choosing appropriate visual encodings

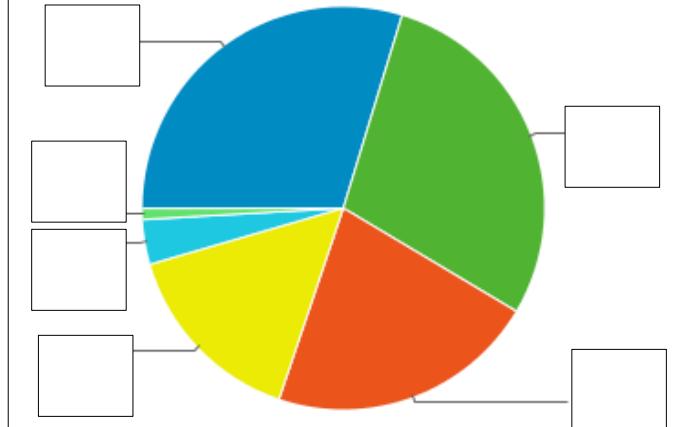
- Use the most easily interpreted encoding for the most important variable
- Respect natural ordering
- Ensure sufficient distinct values

Using Visual Encodings Effectively

These are some things you should think about to help choosing appropriate visual encodings

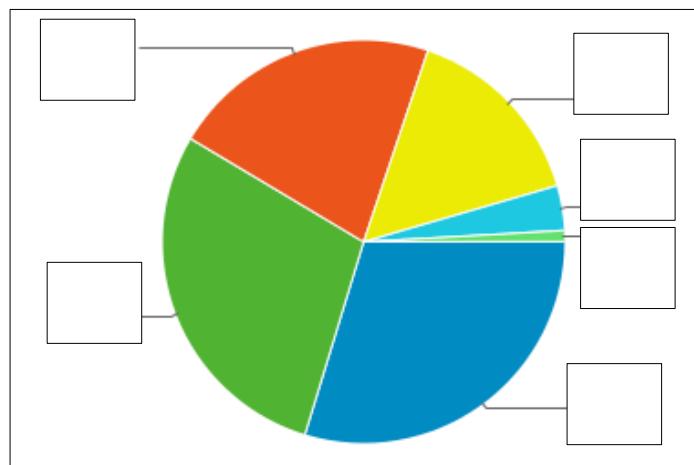
- Use the most easily interpreted encoding for the most important variable**
- Respect natural ordering
- Ensure sufficient distinct values

Angle

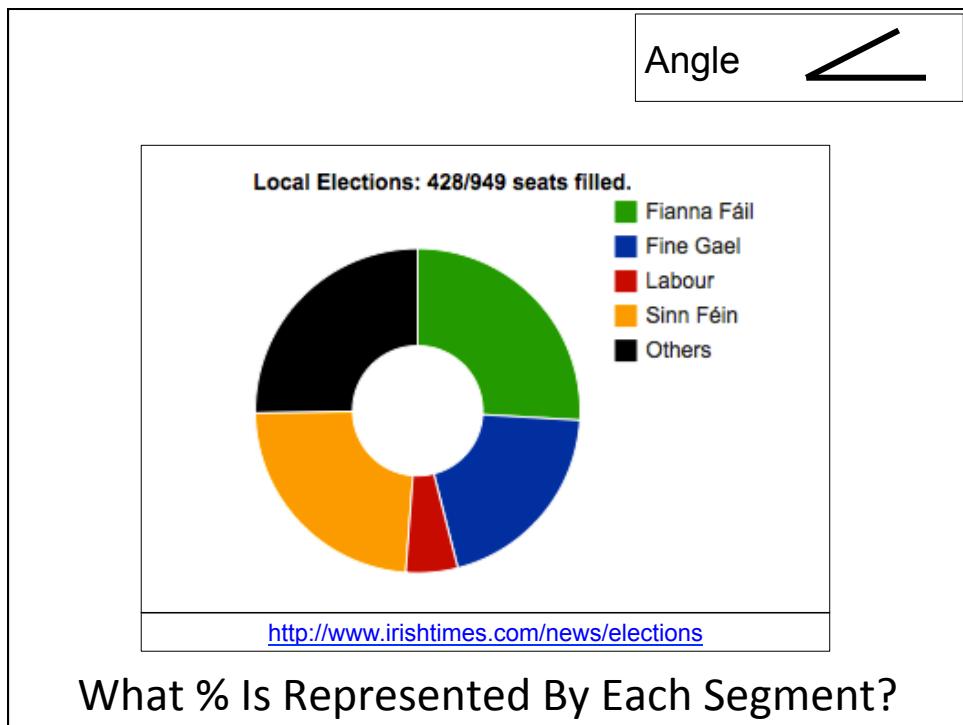
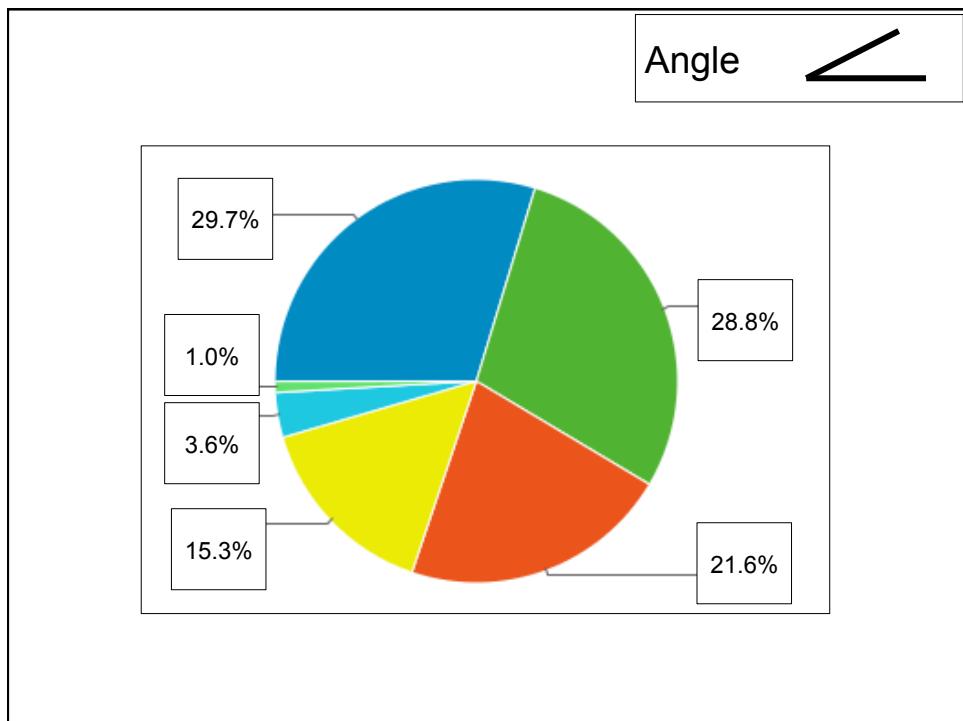


Which Segment Is Biggest?

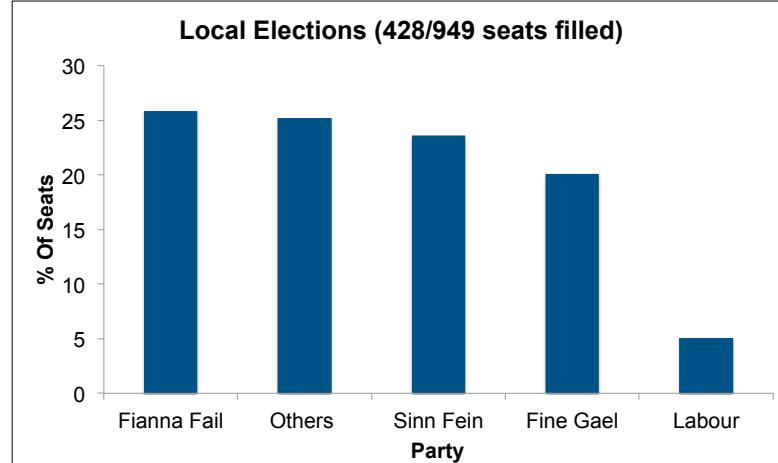
Angle



Which Segment Is Biggest?

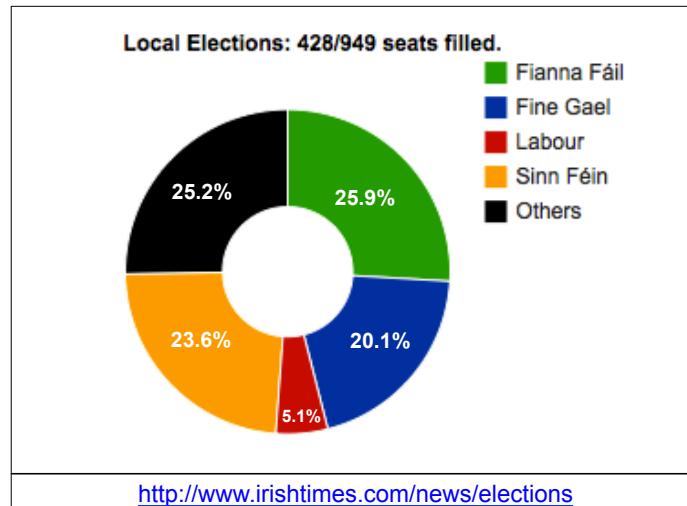


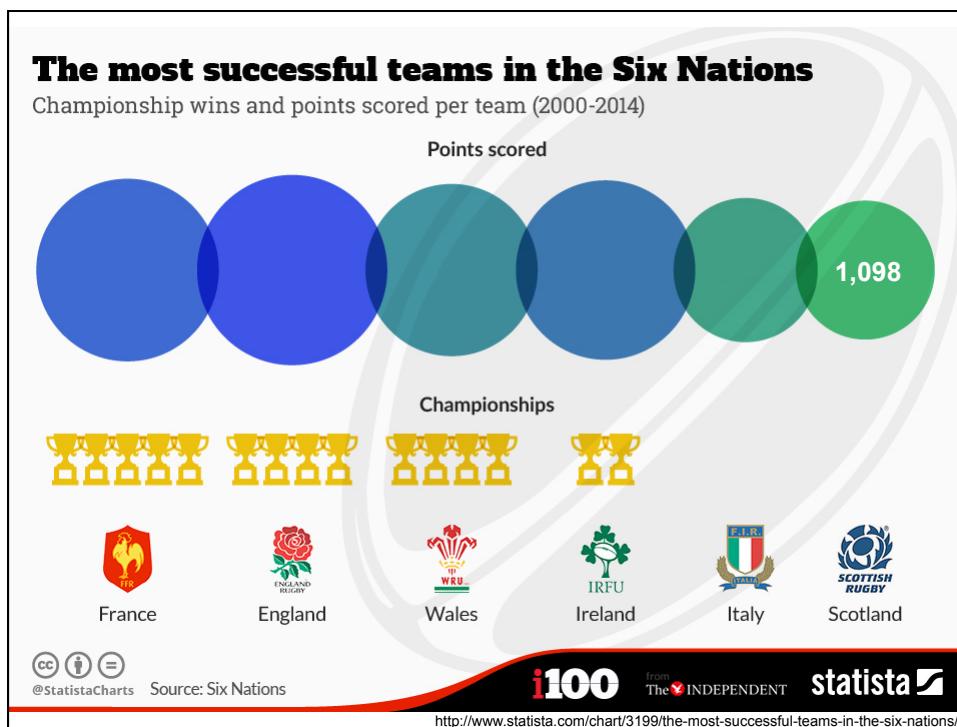
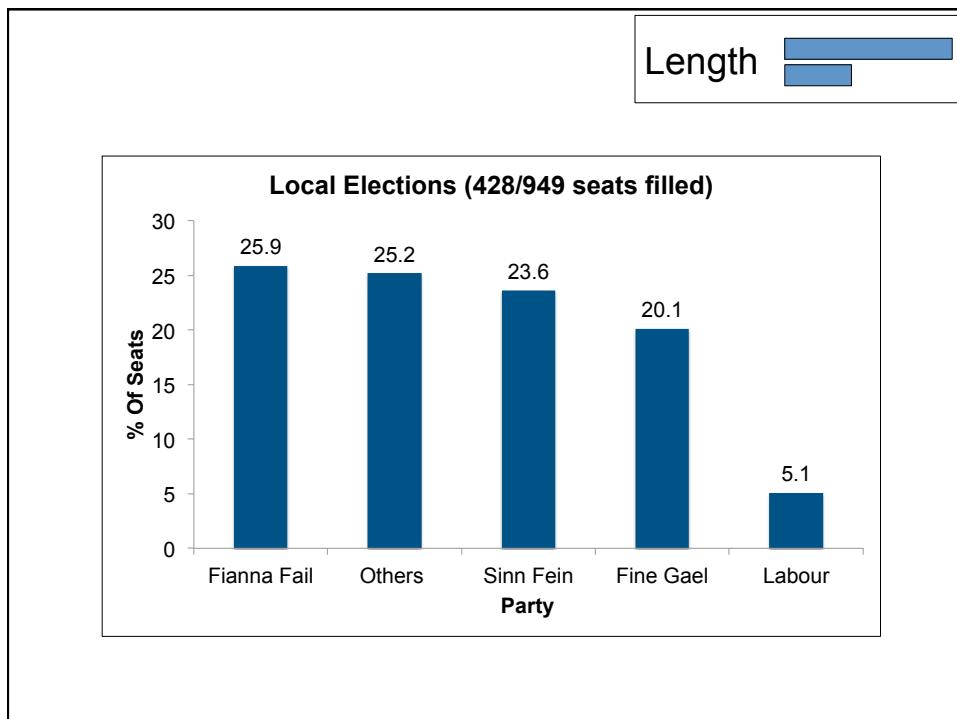
Length

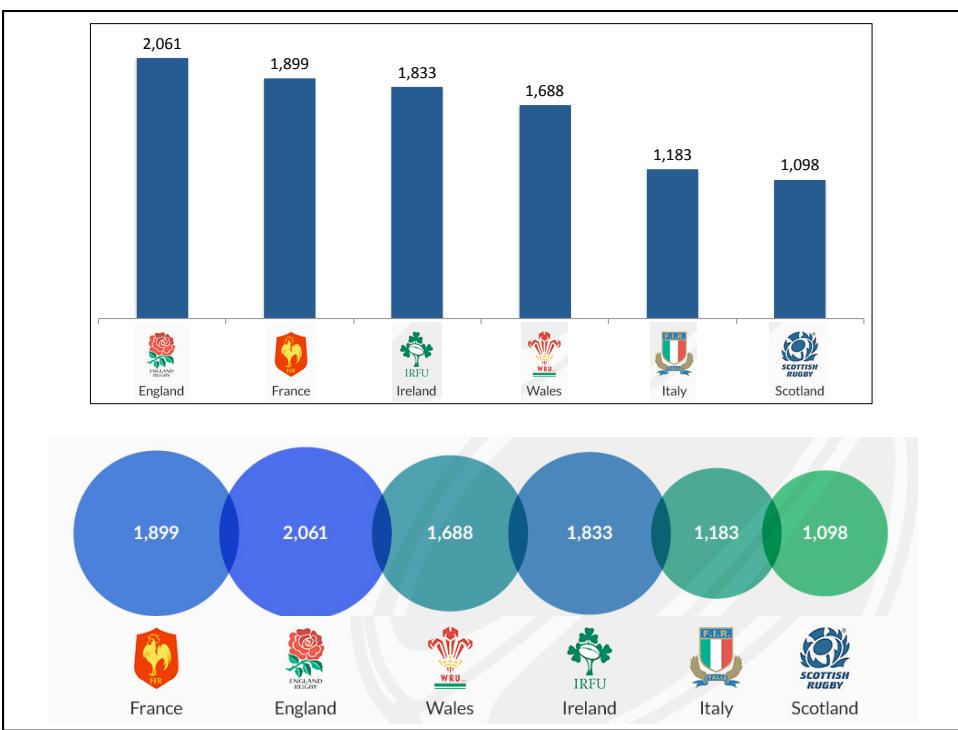
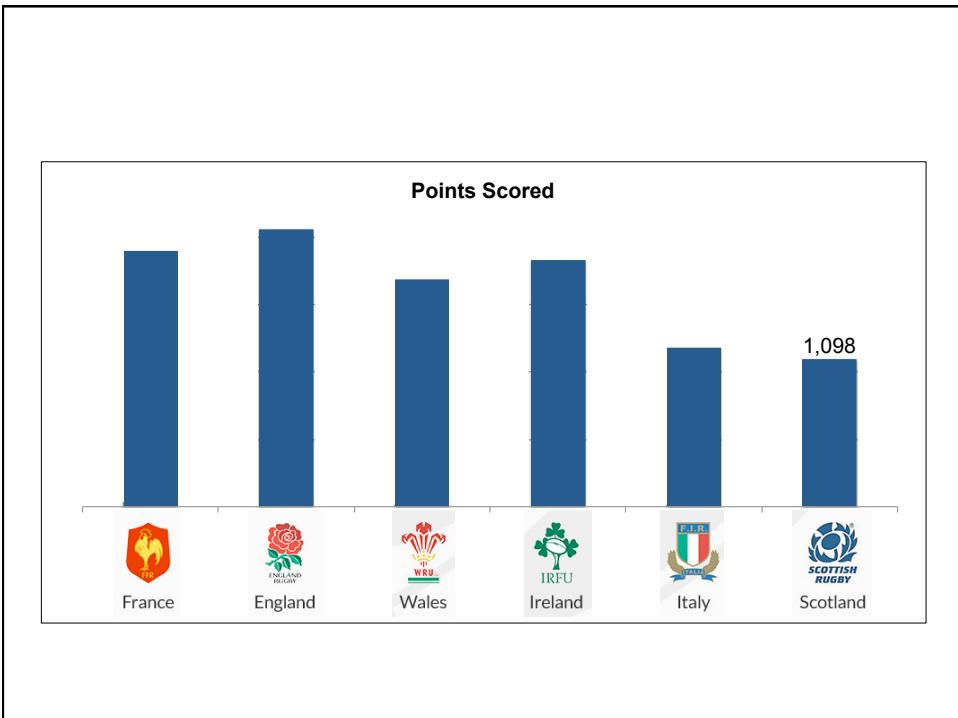


What % Is Represented By Each Segment?

Angle







People Always Get Circle Areas Wrong!



<http://blog.revolutionanalytics.com/2011/01/for-data-visualization-circles-dont-cut-it.html>

People Always Get Circle Areas Wrong!



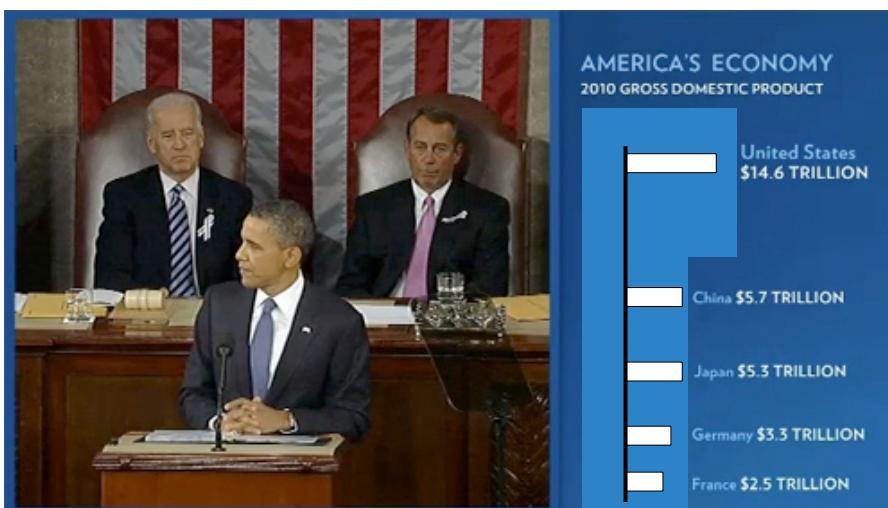
<http://blog.revolutionanalytics.com/2011/01/for-data-visualization-circles-dont-cut-it.html>

People Always Get Circle Areas Wrong!



<http://blog.revolutionanalytics.com/2011/01/for-data-visualization-circles-dont-cut-it.html>

People Always Get Circle Areas Wrong!

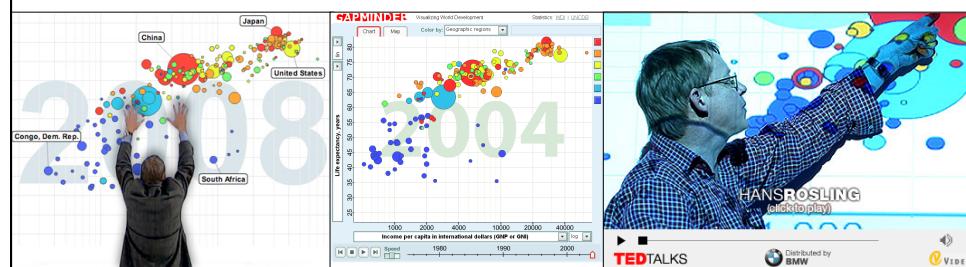


<http://blog.revolutionanalytics.com/2011/01/for-data-visualization-circles-dont-cut-it.html>

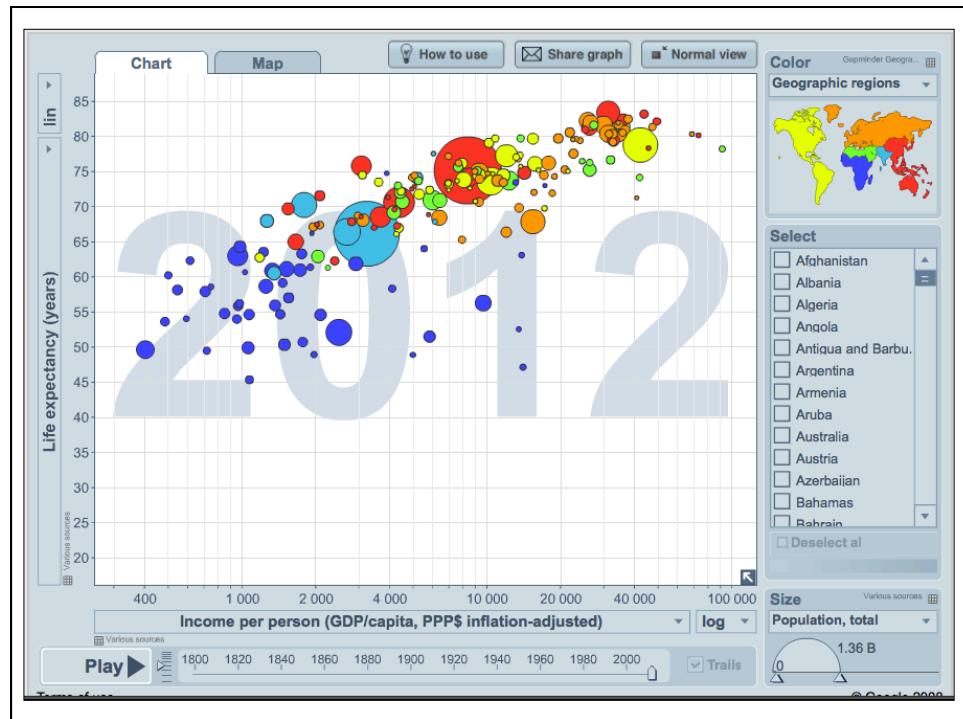
Hans Rosling (Legend!)

Now that we have talked about encodings take a look at this screenshot from Hans Rosling and discuss the encodings used

- How many data dimensions are visually encoded?
- How is importance ordering used?



The tool and data that Rosling uses is available at: www.gapminder.org

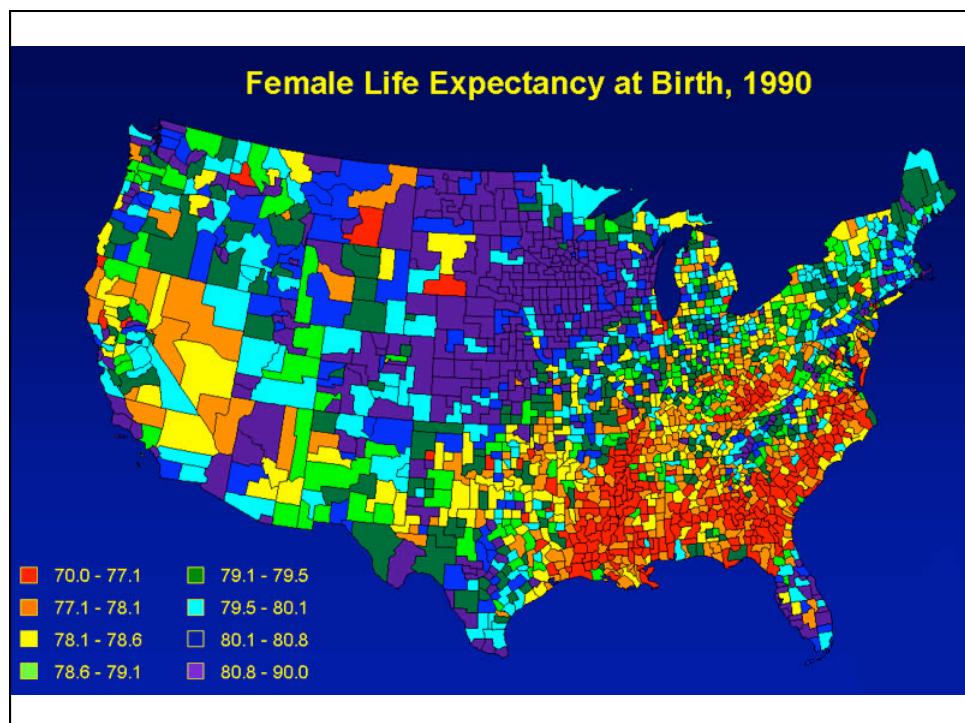
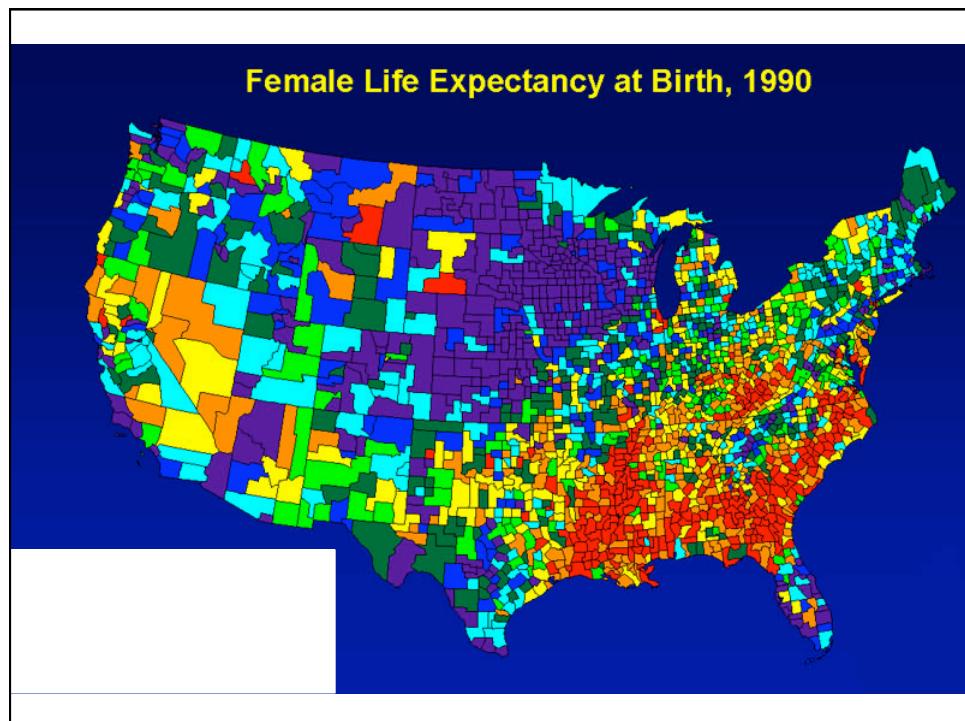


Using Visual Encodings Effectively

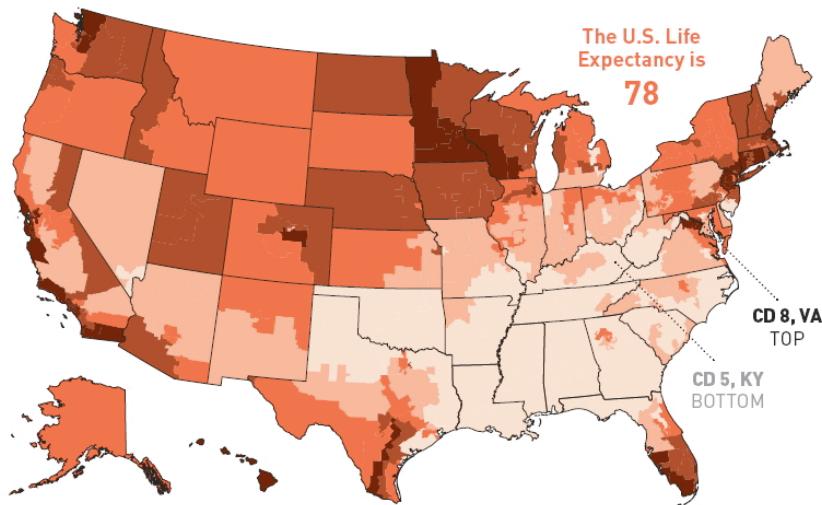
Choosing Appropriate Visual Encodings

- Use the most easily interpreted encoding for the most important variable
- **Respect natural ordering**
- Ensure sufficient distinct values





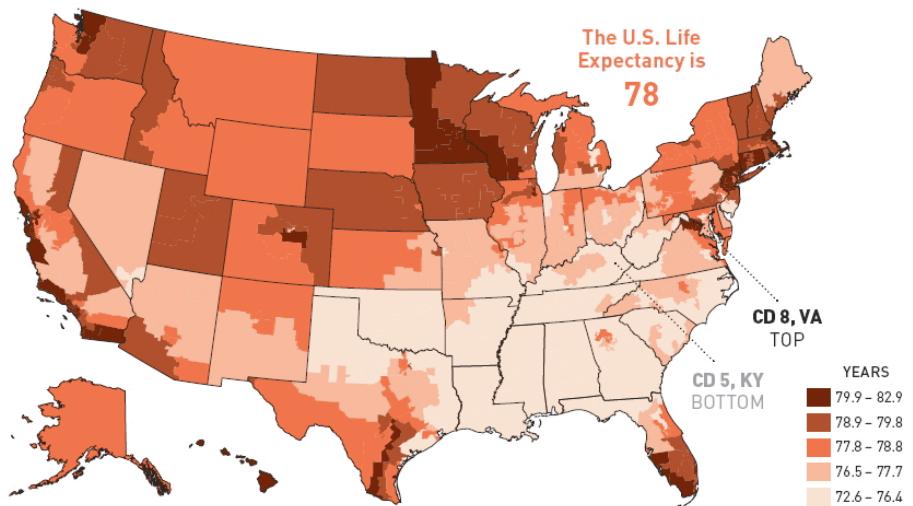
HEALTH INDEX: Life Expectancy at Birth
Great variations in lifespan exist depending on where you live.



www.measureofamerica.org

CONTACT: Alissa Neil PR | 212-431-4411 | MeasureofAmerica@alissaneilpr.com

HEALTH INDEX: Life Expectancy at Birth
Great variations in lifespan exist depending on where you live.



www.measureofamerica.org

CONTACT: Alissa Neil PR | 212-431-4411 | MeasureofAmerica@alissaneilpr.com

Encoding	Example	Ordered	Values
Position		Yes	Many
Length		Yes	Many
Angle		Yes	Moderate
Area		Yes	Moderate
Symbol		No	Few
Colour Hue		No	Few
Colour Value		Yes	Few
Pattern		No	Few

Using Visual Encodings Effectively

Choosing Appropriate Visual Encodings

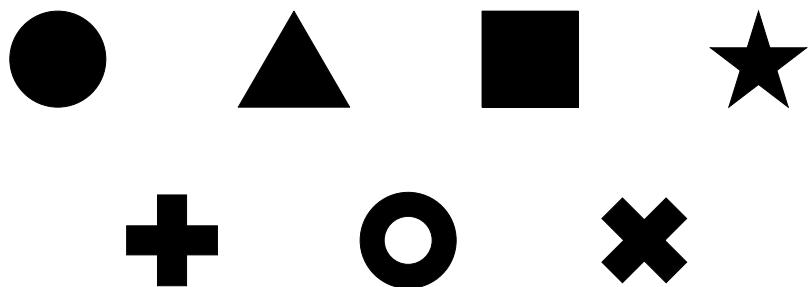
- Use the most easily interpreted encoding for the most important variable
- Respect natural ordering
- **Ensure sufficient distinct values**

Ensure Sufficient Distinct Values

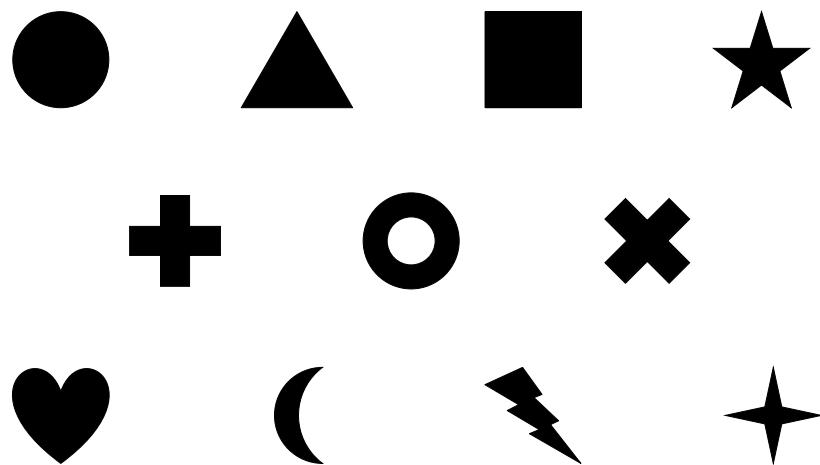
When choosing an encoding for a dimension ensure that the encoding has sufficient capacity to encode all of the values of the data dimension

- For **continuous** data dimensions and infinite number of values is possible and so continues encodings such as position and length work best
- For **categorical** data dimensions there is more flexibility but we can run out of possibilities quickly

You Run Out Of Shapes Very Quickly!



You Run Out Of Shapes Very Quickly!



Distinct Colours

There are a lot of colours in the world, but we cannot tell them apart if too similar

The standard advice for using colour to encode categories is to limit your selection to ideally about 6 - hopefully no more than 12, and absolutely no more than 20 - colours and corresponding categories

There are a range of defined colour palettes that work well

Distinct Colours

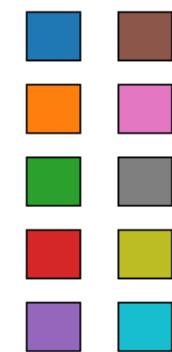
1 Red	
2 Green	
3 Yellow	
4 Blue	
5 Black	
6 White	

7 Pink	
8 Cyan	
9 Gray	
10 Orange	
11 Brown	
12 Purple	

Distinct Colours

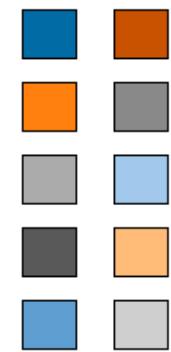
Select Color Palette:

Tableau 10



Select Color Palette:

Color Blind 10



Encoding	Example	Ordered	Values
Position		Yes	Many
Length		Yes	Many
Angle		Yes	Moderate
Area		Yes	Moderate
Symbol		No	Few
Colour Hue		No	Few
Colour Value		Yes	Few
Pattern		No	Few

SUMMARY

Summary

Information visualisation is fundamentally about mapping data dimensions to visual encodings

There are good, well researched guidelines that we can use for doing this - **use them!**