Data Visualisation

Associate Professor Goh Wooi Boon

College of Computing and Data Science
Nanyang Technological University

email: *aswbgoh@ntu.edu.sg*

Chapter 3 – Visual Encoding

Contents

- Visual Encoding
- Deconstructions

Chapter 3.1 – Visual Encoding

Contents

- Visual Variables
- Visual Marks
- Visual Encoding Examples

The Data Visualisation Process

An Overview

- Data visualisation is the transformation of data into an image using an appropriate visual encoding strategies.

Month	Category	Amount
1-Jan	Food	420.30
1-Jan	Transport	200.20
1-Jan	Others	50.00
2-Feb	Food	200.45
2-Feb	Transport	210.00
2-Feb	Others	500.23
3-Mar	Food	150.60
3-Mar	Transport	205.70
3-Mar	Others	80.30

Data

Data & Conceptual Models

Visual Encoding

Mapping from data to image

Image

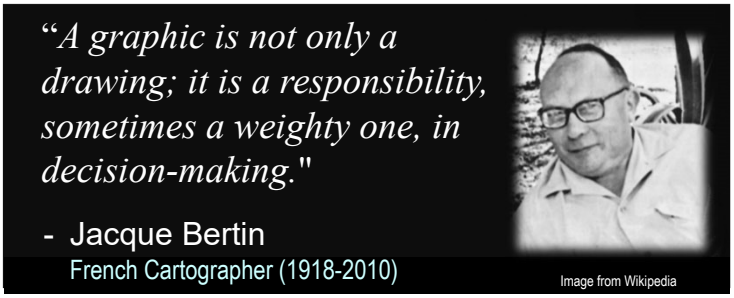
Visual Marks and Channels (Variables)

Lines **Colour**

Visual Variables

Bertin's Visual Variables

- Jacque Bertin, a French cartographer proposed set of **visual variables** in 1967 that can be used to construct map symbols and other graphical techniques.



NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

[1] Bertin's quote from: J. Bertin, graphics and graphic information processing (1981), https://books.google.com.sg/books?id=csqX_xnm4tcC&printsec=copyright&redir_esc=y#v=onepage&q&f=false

5

5

Visual Variables

Bertin's Visual Variables

- Jacque Bertin, a French cartographer proposed set of **visual variables** in 1967 that can be used to construct map symbols and other graphical techniques.
- Bertin identified **seven** main categories of visual variables:
  - position
  - size
  - shape
  - value (intensity)
  - colour
  - orientation
  - texture

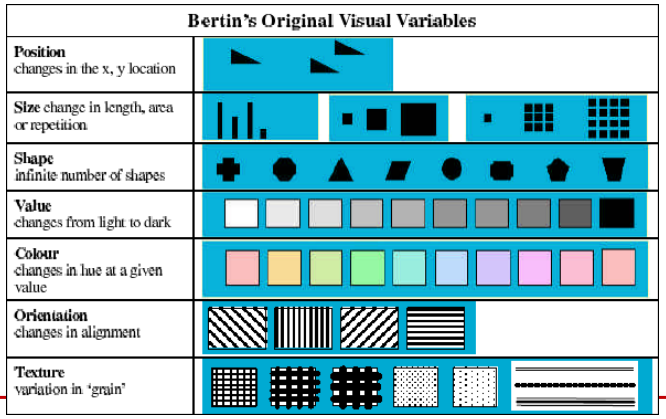


Image from https://infovis-wiki.net/wiki/Visual_Variables

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

6

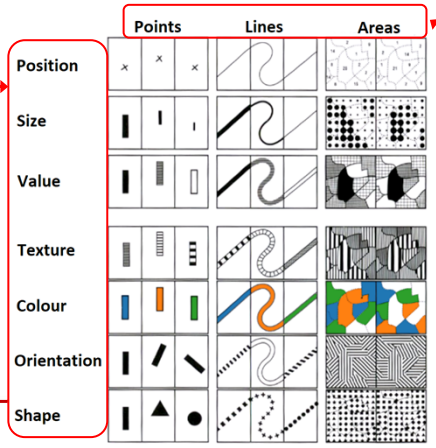
6

Visual Variables

Marks and Channels

- Basic graphic elements (**visual marks**) like points, lines and areas can then be modified with different **perceptual channels** (visual variables) to encode information.

Channels (Visual variables)
Control the appearance of the visual marks



Visual marks
Basic graphical elements in an image

Recap: What's their difference?
Marks are basic geometric elements that depicts items or links. **Channels** control their appearance.

Image adapted from https://graphworkflow.com/retinal/
Bertin's taxonomy of visual encodings from The Semiology of Graphics

Visual Encoding

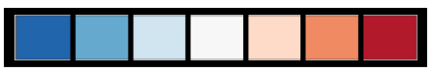
Encoding Effectiveness of Perceptual Channels

- The effectiveness in encoding different **NOIR** data scales depends on the channel.
- **Value** (intensity) is perceived as **ordered** and can encoded **ordinal** scale when suitably **discretised**.

- **Value**, when **continuously varying**, can encode **interval**/**ratio** scales but **lacks precision** due to our poor brightness discrimination ability.

- **Hue** (colour) is normally perceived as **unordered** and can only encode **nominal** scale variables.

- **Hue** and **value,** suitably combined, can encode **ordered** variables with a meaningful **mid-point**.



Visual Variables

Bertin's Levels of Organisation

- Bertin's levels of organisation illustrates the degree of effectiveness in encoding the various **NOIR** data scales.

Nominal
Ordinal
Quantitative

Note 1: **Interval** and **Ratio** is classified as a **Quantitative** scale.

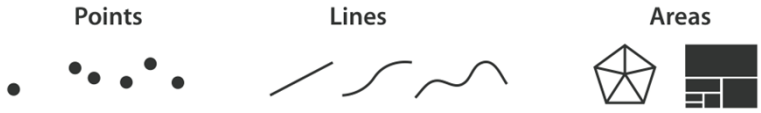
Note 2: **Size** indicates **degree** of encoding effectiveness

Position	**N**	**O**	**Q**
Size	**N**	**O**	**Q**
Value	**N**	**O**	**Q**
Texture	**N**	**O**	
Colour	**N**		
Orientation	**N**		
Shape	**N**		

Visual Marks

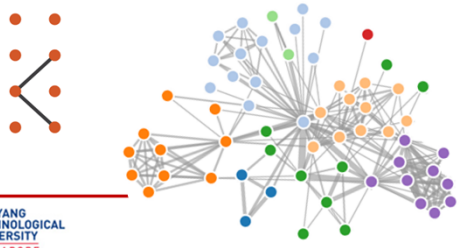
Types of Marks

- **Item marks** – are basic geometric objects that depict **items** on a graph[2].

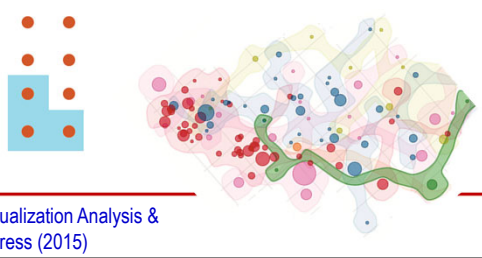


- **Link marks** – show relationship between items[2].

Connection - shows pairwise relationships



Containment - shows group relationships



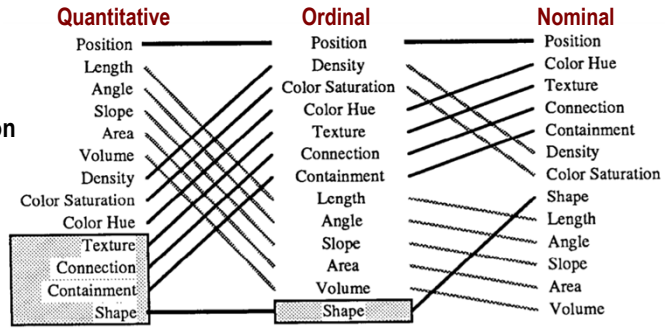
[2] T. Munzner, Visualization Analysis & Design, CRC Press (2015)

Visual Encoding

Mackinlay's Visual Variables

- **Jock Mackinlay**[3] expanded Bertin's variables and proposed an **effectiveness ranking** based on the NOIR data scales.

Note:
Effectiveness is measured by how readily the **information conveyed** by a visualisation is **perceived** by the intended audience.



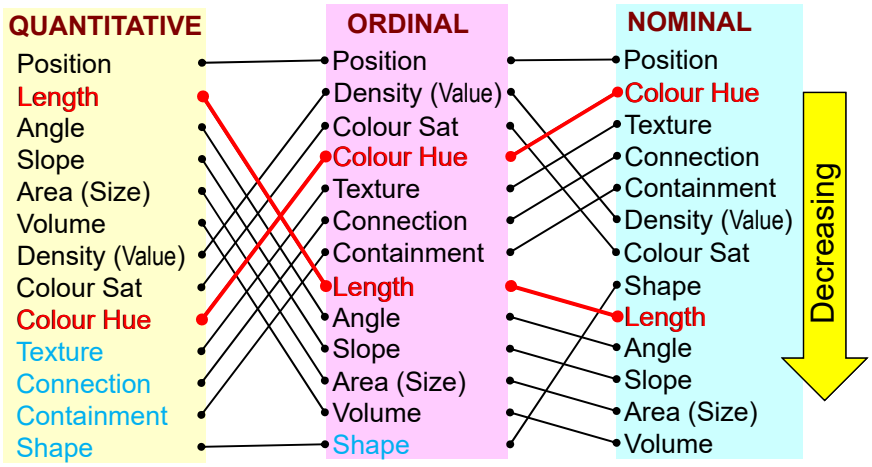
Ranking of perceptual tasks. Tasks shown in the gray boxes are not relevant to these types of data.

Jock D. Mackinlay
Information Visualisation Expert
and Former Vice President of R&D
at Tableau Software
(Image from Twitter)

[3] J. Mackinlay, Automating the Design of Graphical Presentation of Relational Information (1986), https://research.tableau.com/sites/default/files/p110-mackinlay.pdf

Visual Encoding

Mackinlay's Effectiveness Ranking[3]

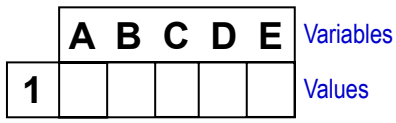


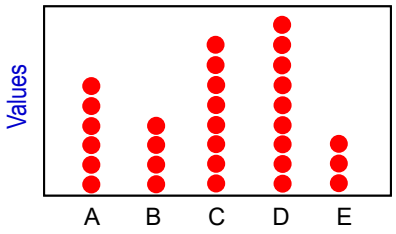
[3] J. Mackinlay, Automating the Design of Graphical Presentation of Relational Information (1986), https://research.tableau.com/sites/default/files/p110-mackinlay.pdf

Visual Encoding Examples

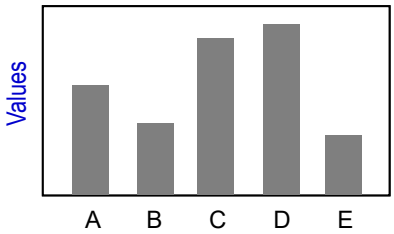
Univariate Data

- In visual encoding of univariate datasets, common visual marks used are **lines** (e.g. line and bar charts) and **points** (e.g. dot plot).





Dot Plot (Points)



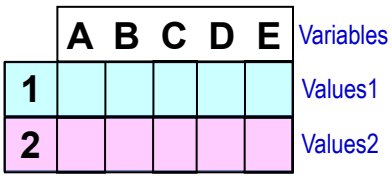
Bar Chart (Lines)

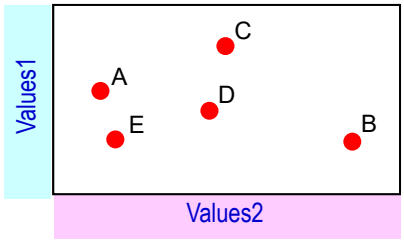
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Visual Encoding Examples

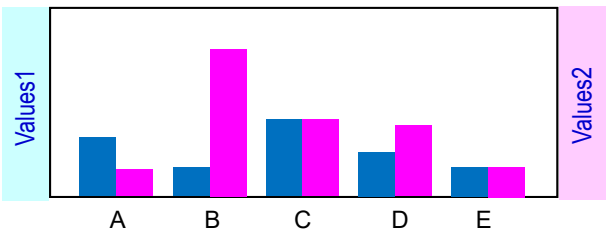
Bivariate Data

- In visual encoding of bivariate datasets, the most common visual mark used are **points** in 2D **scatter plot**. **Line** mark may be used, such as with a **dual Y-axis bar charts**.





2D Scatter Plot (Points)



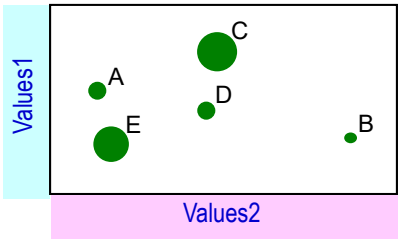
Dual-axis Bar Chart (Lines)

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

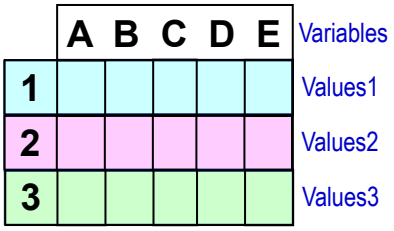
Visual Encoding Examples

Trivariate Data

- Trivariate datasets are commonly encoded using **points**. The two axes of a **2D scatter plot** can encode two quantitative (Q) data type and **area** of the dots can encode the third ordinal or Q data type.



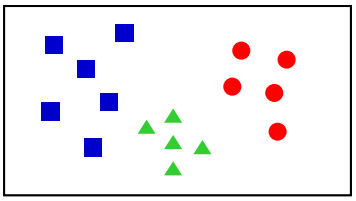
2D Scatter Plot
(with size of dots representing Values3)



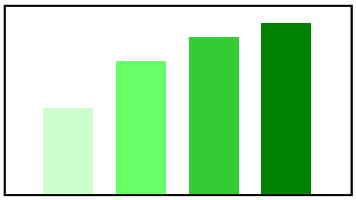
Visual Encoding Examples

Redundant Encoding

- Redundant encoding is the use of **more than one graphical** or visual variable (e.g. color + shape) to encode/represent **one variable of data**.
- Redundancy allows **more perceptual channels** to get the same information into our brain. This helps information to be perceived **faster**, **easier** and more **accurately**[4]. It should be considered if there are leftover unused visual variables.



2D Scatter Plot Example
(Redundant **Colour** + **Shape**)



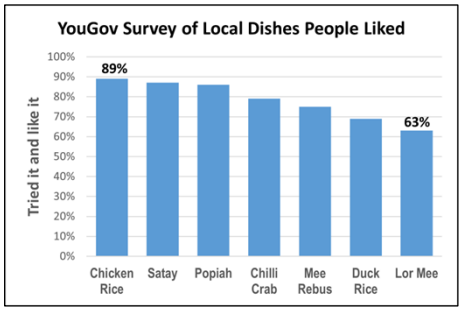
Bar Chart Example
(Redundant **Length** + **Value**)

[4] C. Ware, Information Visualization: Perception for Design, 3rd Edition, Morgan Kaufmann (2013),

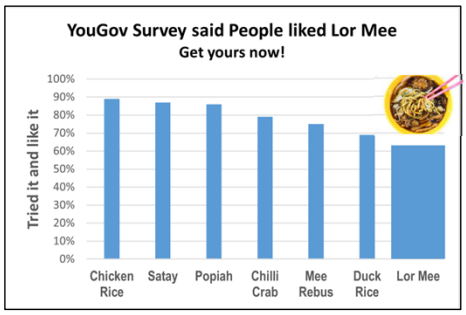
Ethical Visualisation

Use fair and consistent visual encoding

- Your stall specialises in Lor Mee and you saw the results of the recent YouGov local food survey. You decided to share these results with your potential customers.



The YouGov survey results of selected local dishes



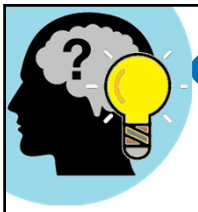
Your visualisation of the YouGov food survey

Has ethical visualisation been violated here?

Note: Height and width are integral perceptual dimensions. People tend to see the size of the bar as representing the data value [5].

Note: Food survey figures from: https://sg.yougov.com/en-sg/news/2019/08/13/classic-singaporean-cuisine-ranked-singaporeans/

[5] Stephen M. Kosslyn, Graph Design for the Eye and Mind, Oxford University Press (2006)



Think and Apply

Visual Encoding Exercise

- Sketch a possible chart that uses an appropriate visual encoding design for each of the data tables. Consider the given data attributes carefully in your encoding design.

Living Members from Class of '45

	Names	Gender	Age
1			
2			
3			
4			
5			
6			

Visual mark – **lines**

Orientation T-shirts - Comparing Halls

	No.	Gender	Weight	Height	Hall (A,B,C,D)	T-shirt Size
1						
2						
3						
:						
:						
700						

Visual mark – **points**

COVID-19 Infection in 2020

	Dates	Total Cases	Warded	ICU
1				
2				
3				
:				
:				
365				

Visual mark – **areas**

Summary

Visual Encoding

- Data visualisation is the transformation of **data** into an **image** that is **informative** and conveys **useful insights** embedded within the data.
- **Visual encoding** is the process of **mapping** these data into that image.
- Based on the attributes of the various dimensions of the dataset, appropriate combinations of visual **marks** and **variables** (channels) are used to encode the data.
- Effective visual encoding requires a good understanding of the characteristics and limitations of **human visual perception**.

Chapter 3.2 – Deconstruction

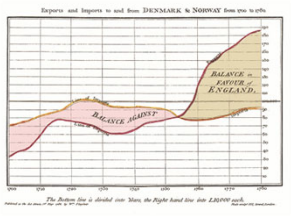
Contents

- Playfair's Import/Export (Line Charts)
- Wattenberg's Map of the Market (Tree Map)
- Fry's Basketball Salary versus Performance (Slope Chart)

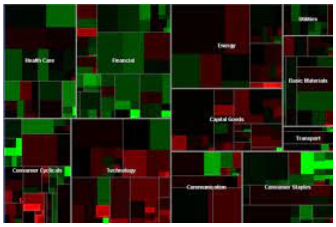
Deconstruction

Seeing How Others Do It

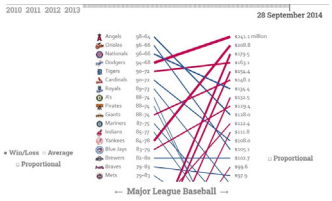
- The process of deconstruction allows us to **analyse** how others have visually encoded the information they wanted to convey in their dataset.
- Deconstruction is a great way to **learn how to construct** your own visualisation, especially if the examples analysed have been effective in telling their visual story.



Playfair's Import/Export (1786)
A good old pre-digital chart



Wattenberg's Map of the Market (1998)
A visualisation still current in use
checkout: https://finviz.com/map.ashx



Fry's Salary vs Performance (2014)
Visualisation for US Basketball fans
checkout: https://fathom.info/salaryper/

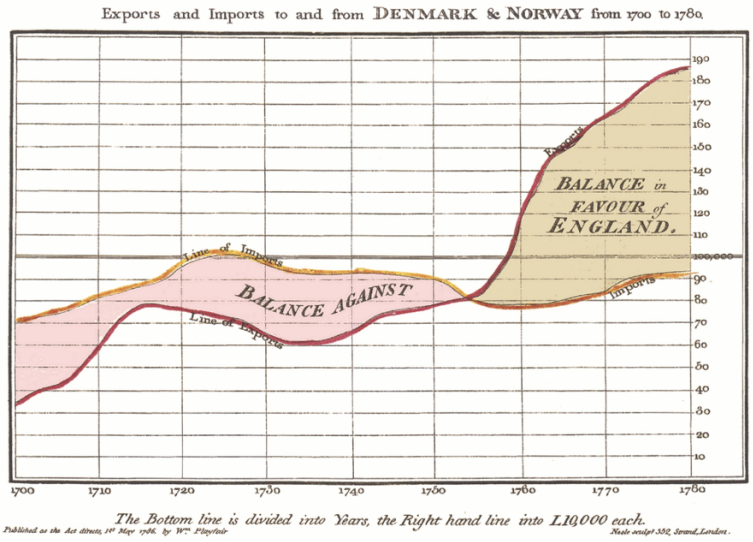
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

21

21

Deconstruction - Example #1

William Playfair's Line Charts (1786)



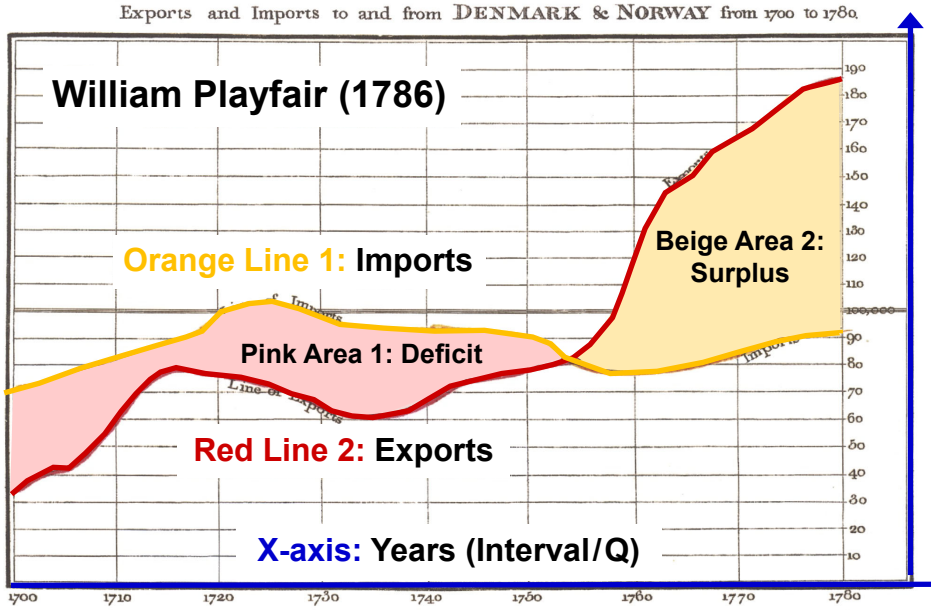
- William Playfair's Commercial and Political Atlas was published in 1786
- Playfair examined the imports and exports between Britain and various countries.
- To illustrate these trade relationships, Playfair created the first ever line charts that show the change in trade over time.

[6] Image can be found at: https://upload.wikimedia.org/wikipedia/commons/5/52/Playfair_TimeSeries-2.png

22

22

Deconstruction - Example #1



Visual Variables

Lines

Q 1. Position (X,Y)

N 2. Colour (Orange/Red) Import & Export

Areas

N 1. Colour (Pink/Beige) Deficit & Surplus

Y-axis: Currency (£10k) **(Ratio/Q)**

23

23

Deconstruction - Example #2

Martin Wattenberg's Map of the Market (1998)

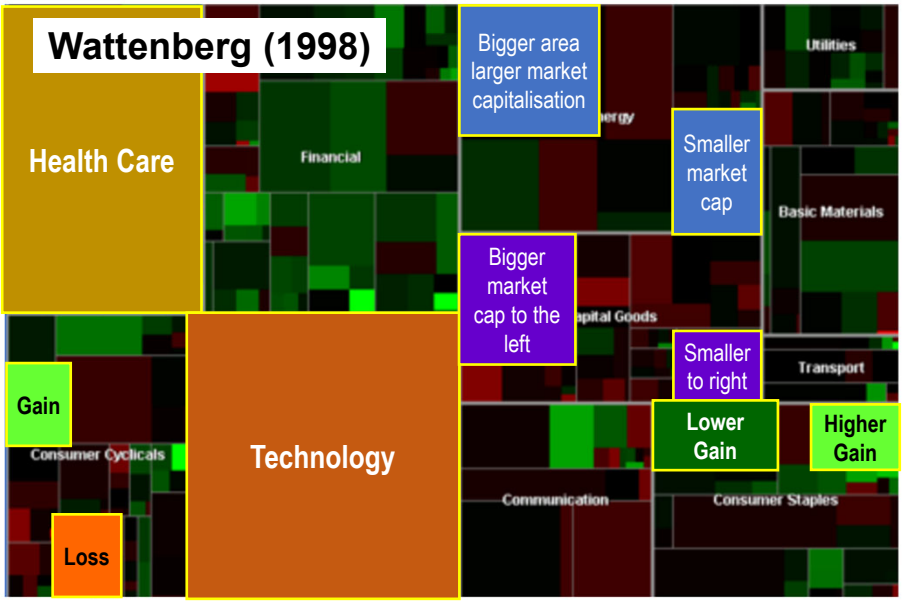
- Martin Wattenberg & co. created one of the 1st web-based visualisations, Map of the Market.
- It displayed live stock market data for more than a decade, starting in 1998, while Martin was at SmartMoney.com.
- Map of the Market uses a modified **Tree Map** algo that creates squarish tiles.

[7] Map of the Market details can be found at: http://www.bewitched.com/marketmap.html

24

24

Deconstruction - Example #2



Visual Variables

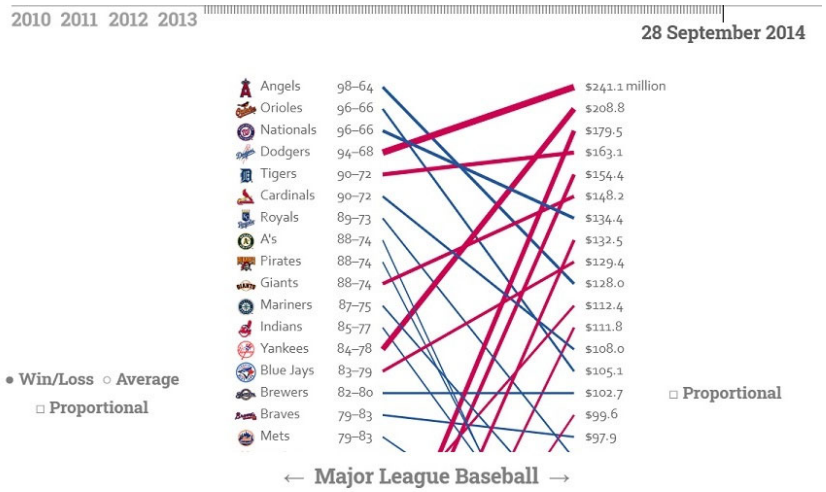
Areas (Rectangle)

- **Q** 1. Size (Market Cap)
- **N** 2. Position (Market Sector)
- **Q** 3. Position (Market Cap)
- **N** 4. Colour Hue
  - (Green - gain)
  - (Red – loss)
- **Q** 5. Colour Saturation
  - (Bright – higher gain)
  - (Dark – lower gain)
  - (Bright – higher loss)
  - (Dark – lower loss)

25

Deconstruction - Example #3

Ben Fry's Slope Chart (2010-2014)

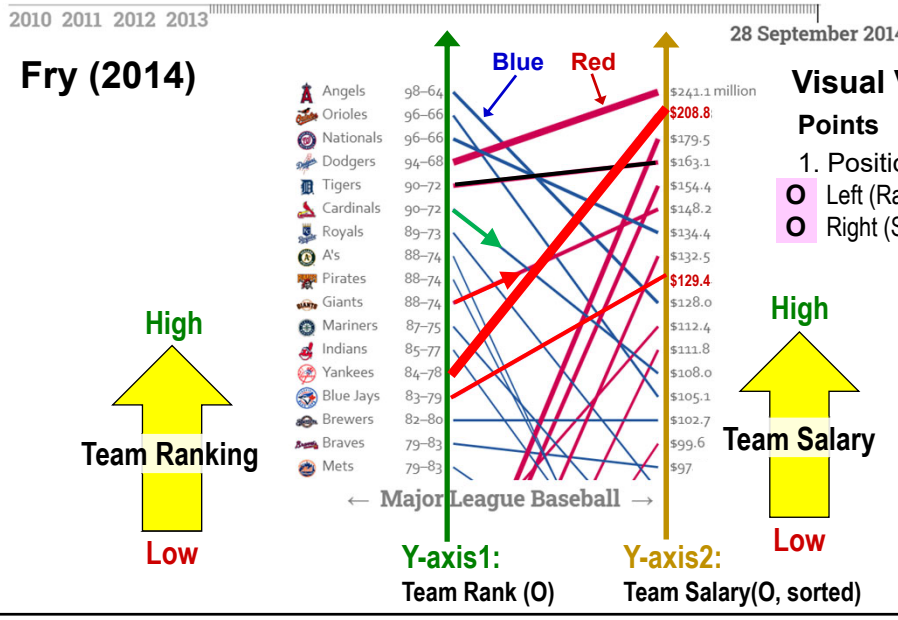


- Ben Fry is a visualisation expert and principal of Fathom, a design and software consultancy in Boston, USA.
- He designed Salary vs. Performance to visualised if baseball teams are spending their money well over the season.
- The visualisation uses a slope chart.

[8] Fry's website for Salary vs Performance: https://benfry.com/salaryper/

26

Deconstruction - Example #3



Visual Variables

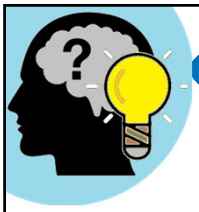
Points

1. Position (Y1, Y2)
   - O Left (Ranking)
   - O Right (Salaries)

Links

2. Position (Team - Salary)
3. Slope
   - O ↘ (Good value)
   - O ↗ (Poor value)
4. Colour Hue
   - N (Blue +ve value)
   - (Red -ve value)
5. Size (thickness)
   - Q (Salary)

27



Think and Apply

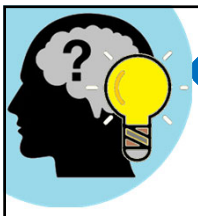
Visualising Relationships

- Joshua Keating & Chris Kirk (@Slate.com) created the **Middle East Friendship Chart** to visualise the complex relationships between different players in Syria's civil war.

- Draw the table for the relational data model.
- What visual marks and variables are used here?
- Redundant encoding?
- Colour choices?
- Effective?

[9] Find Middle East Friendship Chart at: http://www.slate.com/blogs/the_slatest/2015/10/06/syrian_conflict_relationships_explained.html

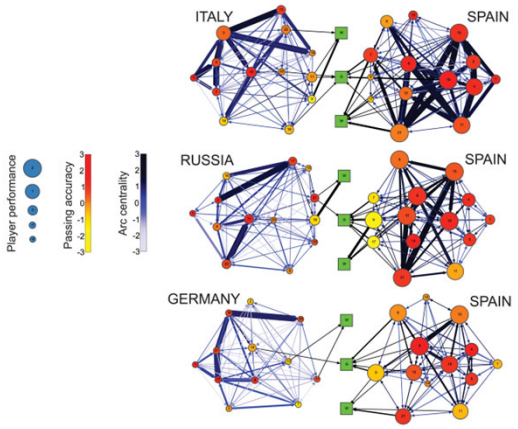
28



Think and Apply

Visualising Performance in Team Games (optional)

- Duch et al. analysed three Euro 2008 matches of the Spanish team using a visualisation technique from social network analysis.



- What types of data variables are visually represented in the image.
- What perceptual channels were used to encode their values?
- Colour choices? Can be better?

[10] Visualisation described in this paper: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0010937

29

29

Summary

Deconstruction

- Data visualisation is both an **art** and a **science**.
- Designing **informative** and **visually appealing** visualisations requires us to learn from many other visualisation experts.
- We can do this by **deconstructing** their good pieces of work.
- We observe how they make use of the various visual marks and variables to **encode useful information** embedded in the datasets.
- We need to do these observation with a **critical eye,** to see what has **worked well** and what has not and **can be improved**.

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

30

30

References for Visual Encoding and Deconstruction

[1] Bertin's quote from: J. Bertin, graphics and graphic information processing (1981) - https://books.google.com.sg/books?id=csqX_xnm4tcC&printsec=copyright&redir_esc=y#v=onepage&q&f=false

[2] T. Munzner, Visualization Analysis & Design, CRC Press (2015)

[3] J. Mackinlay, Automating the Design of Graphical Presentation of Relational Information (1986), https://research.tableau.com/sites/default/files/p110-mackinlay.pdf

[4] C. Ware, Information Visualization: Perception for Design, 3rd Edition, Morgan Kaufmann (2013),

[5] William Playfair's image can be found at: https://upload.wikimedia.org/wikipedia/commons/5/52/Playfair_TimeSeries-2.png

[6] William Playfair's image can be found at: https://upload.wikimedia.org/wikipedia/commons/5/52/Playfair_TimeSeries-2.png

[7] Martin Wattenberg's Map of the Market details can be found at: http://www.bewitched.com/marketmap.html

[8] Ben Fry's website for Salary vs Performance: https://benfry.com/salaryper

[9] Keating & Kirk's Middle East Friendship Chart at: http://www.slate.com/blogs/the_slatest/2015/10/06/syrian_conflict_relationships_explained.html

[10] Duch et al.'s visualisation described in this paper: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0010937

Note: All online articles were accessed between May to June 2021