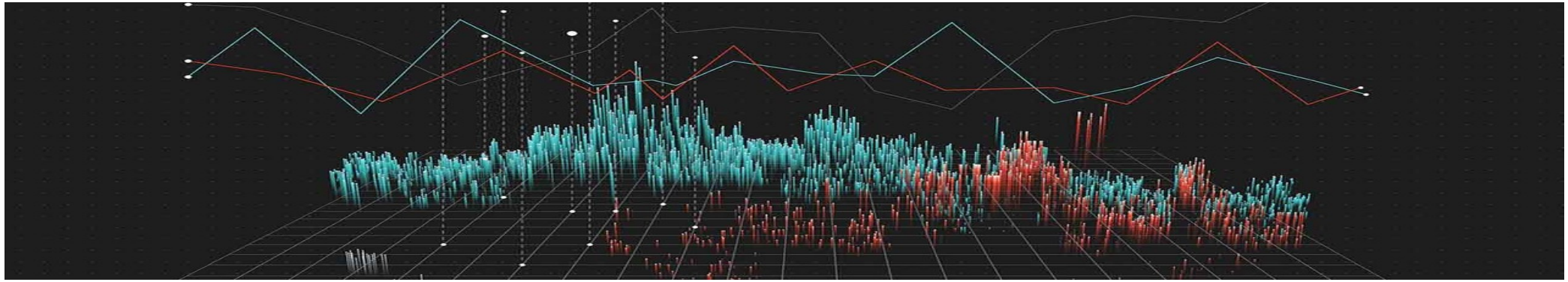
Visualising Data
Lecture 1 - Introduction



Overview

- Preliminaries
- Visualisation
- Module Overview
- Books
- Tools
- Video -The Joy of Stats

PRELIMINARIES

My Details

- Name: Patricia O'Byrne
- Email: patricia.obyrne@tudublin.ie

- Brightspace (Notes, Labs, Assignments):
  - Visualising Data CMPU4091: 2020-21
- Socrative room:
  - POBYRNE (Download from socrative.com/apps/)

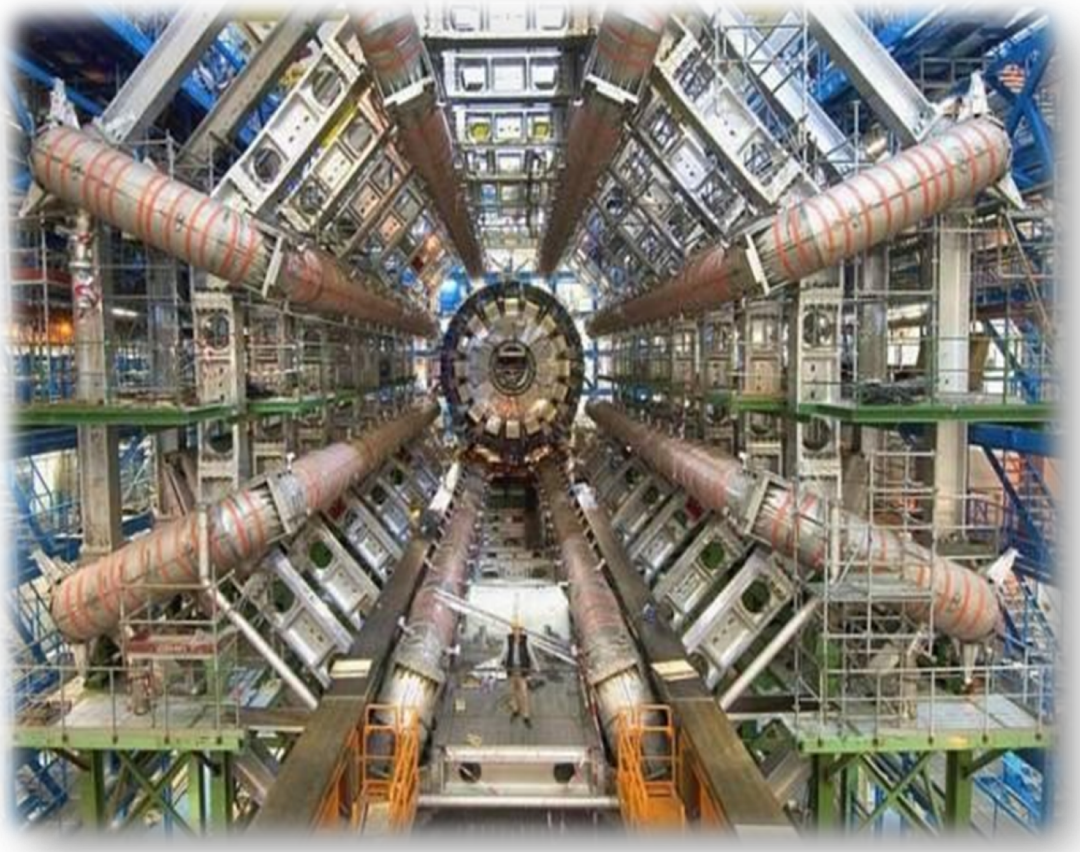
Course Details

- Weekly Labs (24%)
  - 8 x 3% *Arranged to avoid peak project load.*
- Assignment (26%): Visualisation Project
- Exam (50%)

VISUALISATION

Why? - Drowning in Data

- The Large Hadron Collider at CERN is the largest scientific instrument ever constructed
- 1GB of data per second
- 15 PB per year
  - 1 PB = 1 million GB = 1000 TB



Why? - Drowning in Data

- By 2019
  - Video
    - 500 hours of video uploaded per minute
    - 5+ billion videos watched per day
    - 5 billion videos shared to date
  - Users
    - 1.9 billion log-in monthly users
    - Most viewed video 5 billion views
    - First 24 hours 43.2 million views

https://expandedramblings.com/index.php/youtube-statistics/
https://www.omnicoreagency.com/youtube-statistics/

IoT data: Taxi!

- I order a taxi from the airport
- How can this journey be logged?

Data transfer example: Taxi!

- The taxi service operator can record the time the taxi was ordered, the pick-up and destination locations, the customer's details
- The phone company can record the phone call maker and receiver, the duration of the call
- The credit card processing software can record the transaction time, type and amount
- The credit card issuer can record the billing record, the interest rate and available credit

IoT data: Taxi!

- The car maintenance service can record the fuel usage, the engine state
- The passenger can record the expense of the taxi
- The DAA CCTV records the time the taxi drop off, the number of passengers in the car
- Dublin City Council can record the passage of the taxi across the city
- The taxi driver can record the time of the fare, the type of fare, the amount of the fare
- The taxi company can record the path taken by the driver, the duration of the trip

Questions?

- So how can we understand this data?
- So how can we learn for this data?
- How can we make this data accessible?
- How can we gain insights into this data?

What is Visualisation?

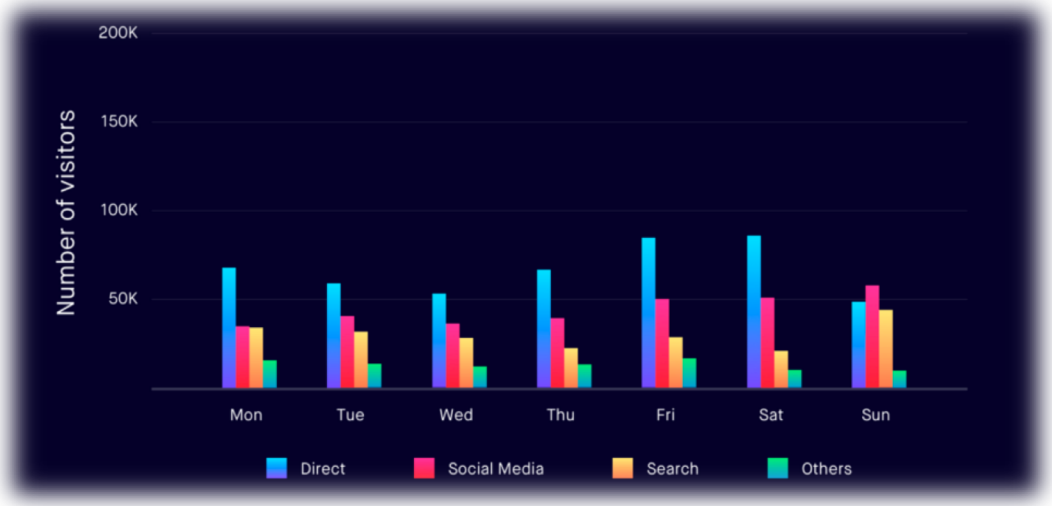
- “Transformation of the symbolic into the geometric”
  - [McCormick et al. 1987]

- “... finding the artificial memory that best supports our natural means of perception.”
  - [Bertin 1967]

- “The use of computer-generated, interactive, visual representations of data to amplify cognition.”
  - [Card, Mackinlay, & Shneiderman 1999]

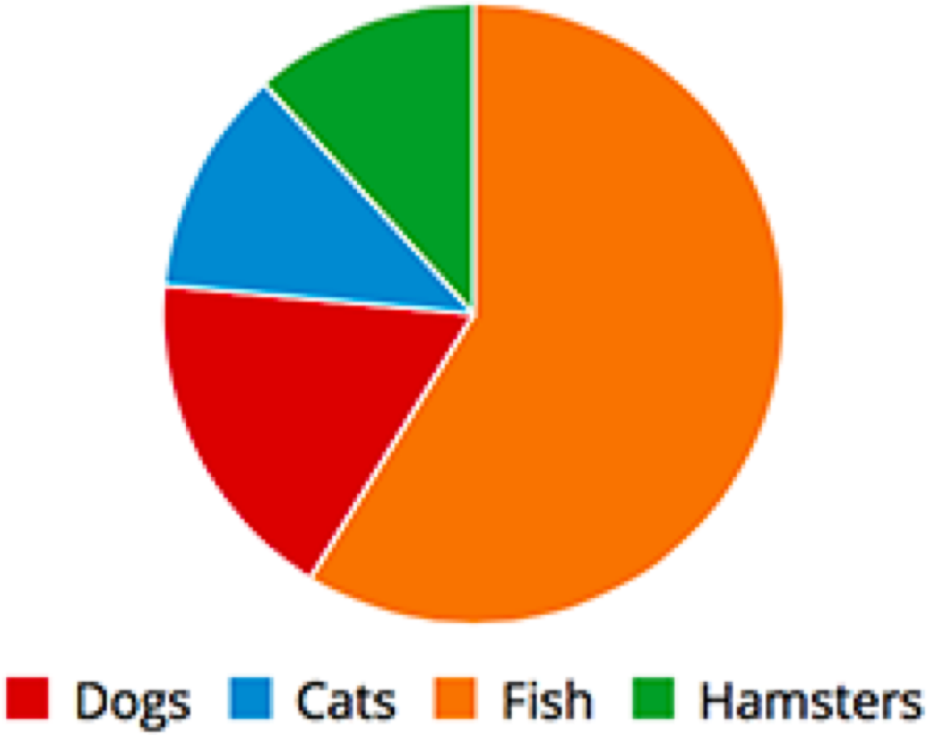
Why Visualisation?

- Answer questions (or discover them)
- Make decisions
- See data in context
- Expand memory
- Support graphical calculation
- Find patterns
- Present argument or tell a story
- Inspire



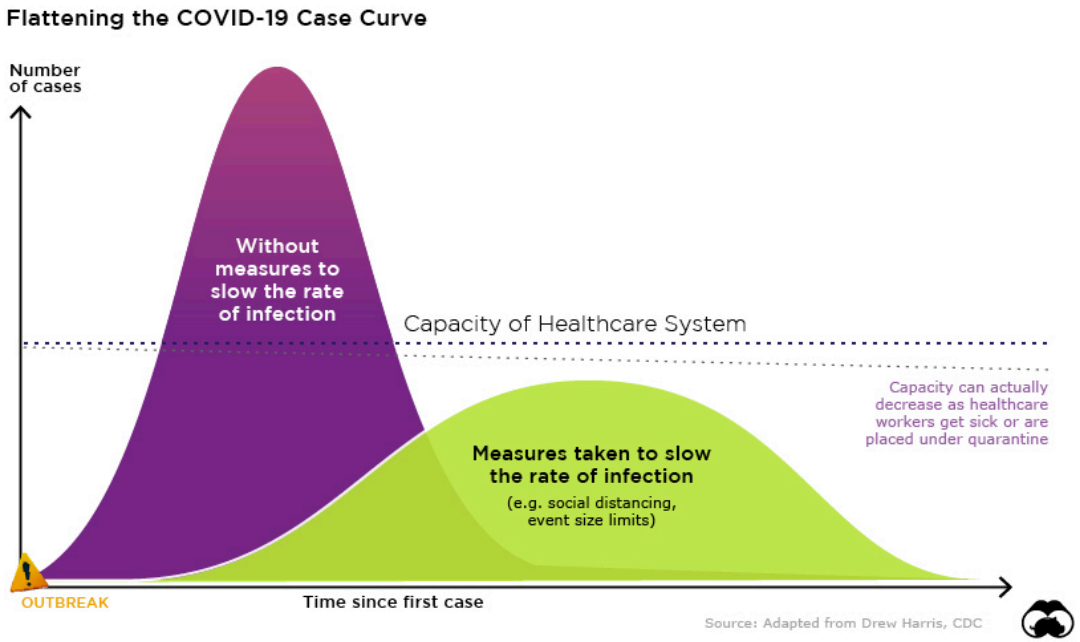
Why Visualisation?

- Answer questions (or discover them)
- Make decisions
- See data in context
- Expand memory
- Support graphical calculation
- Find patterns
- Present argument or tell a story
- Inspire



Why Visualisation?

- Answer questions (or discover them)
- Make decisions
- See data in context
- Expand memory
- Support graphical calculation
- Find patterns
- Present argument or tell a story
- Inspire



Why Visualisation?

Well-designed visual representations can replace cognitive calculations with simple perceptual inferences and improve comprehension, memory, and decision making

By making data more accessible and appealing, visual representations may also help engage more diverse audiences in exploration and analysis

The challenge is to create effective and engaging visualizations that are appropriate to the data

"A Tour Through the Visualization Zoo", J. Heer, M. Bostock, & V. Ogievetsky, Communications of the ACM, vol. 53, no. 6, 2010
http://queue.acm.org/detail.cfm?id=1805128

Why Visualisation?

Well-designed visual representations can replace cognitive calculations with simple perceptual inferences and improve comprehension, memory, and decision making

By making data more accessible and appealing, visual representations may also help engage more diverse audiences in exploration and analysis

The challenge is to create effective and engaging visualizations that are appropriate to the data

"A Tour Through the Visualization Zoo", J. Heer, M. Bostock, & V. Ogievetsky, Communications of the ACM, vol. 53, no. 6, 2010
http://queue.acm.org/detail.cfm?id=1805128

Why Visualisation?

Well-designed visual representations can replace cognitive calculations with simple perceptual inferences and improve comprehension, memory, and decision making

By making data more accessible and appealing, visual representations may also help engage more diverse audiences in exploration and analysis

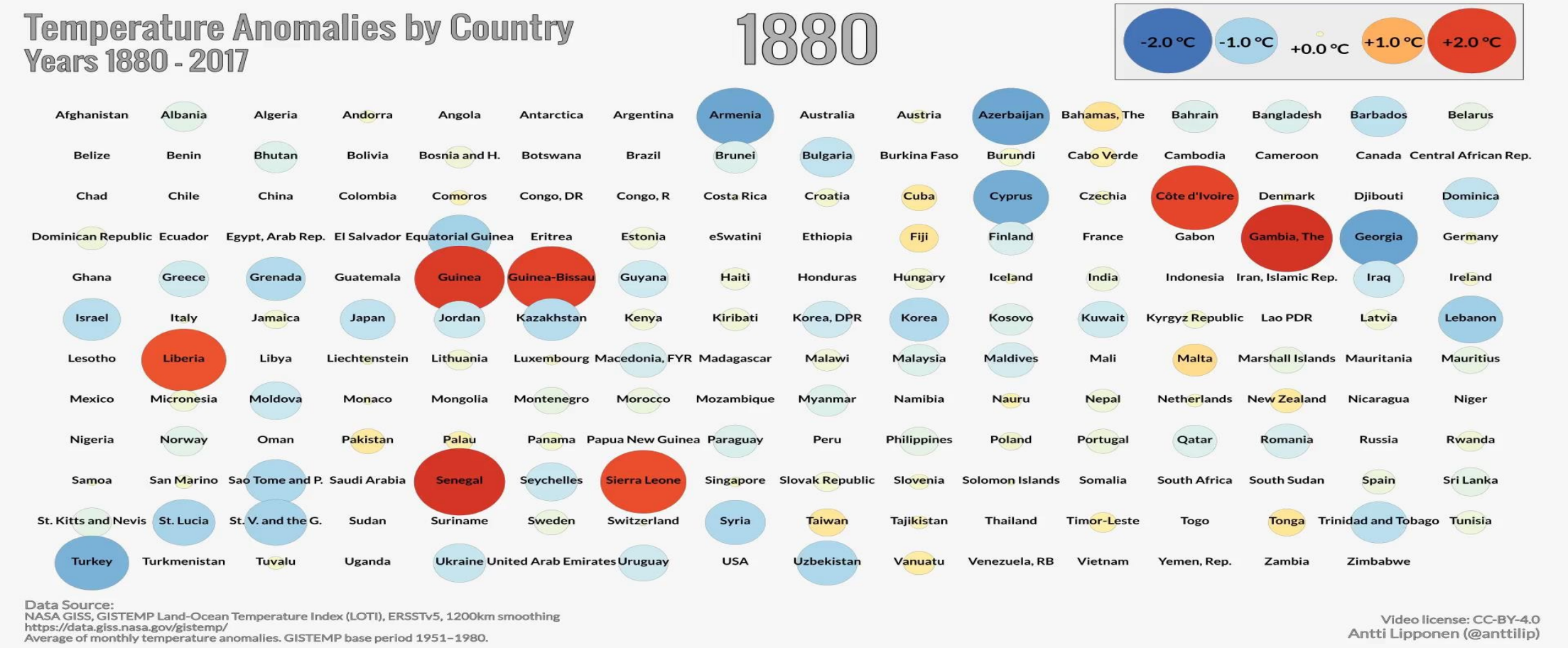
The challenge is to create effective and engaging visualizations that are appropriate to the data

"A Tour Through the Visualization Zoo", J. Heer, M. Bostock, & V. Ogievetsky, Communications of the ACM, vol. 53, no. 6, 2010
http://queue.acm.org/detail.cfm?id=1805128

Functions of Visualisations

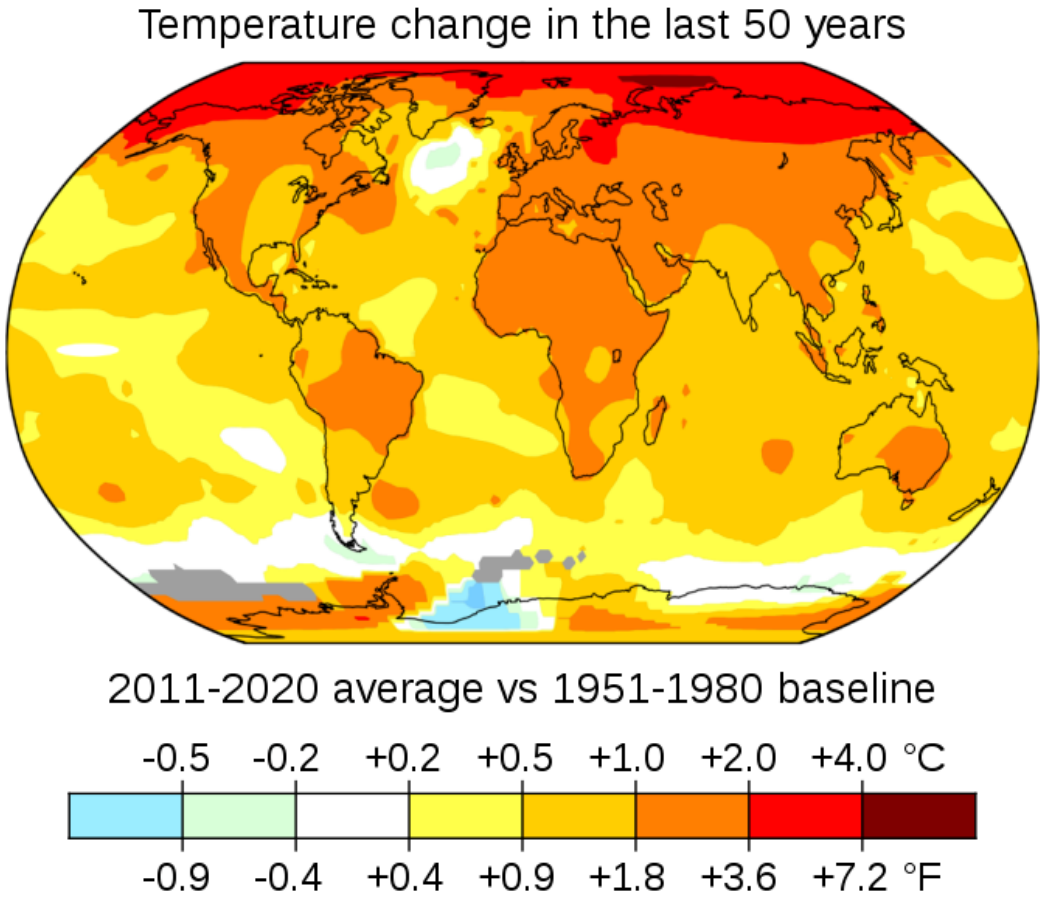
- **Record**: store information
  - Photographs, blueprints, …
- **Analyze**: support reasoning about information
  - Process and calculate
  - Reason about data
  - Feedback and interaction
- **Communicate**: convey information to others
  - Share and persuade
  - Collaborate and revise
  - Emphasize important aspects of data

What is a Visualisation?



https://www.flickr.com/photos/150411108@N06/43350961005/

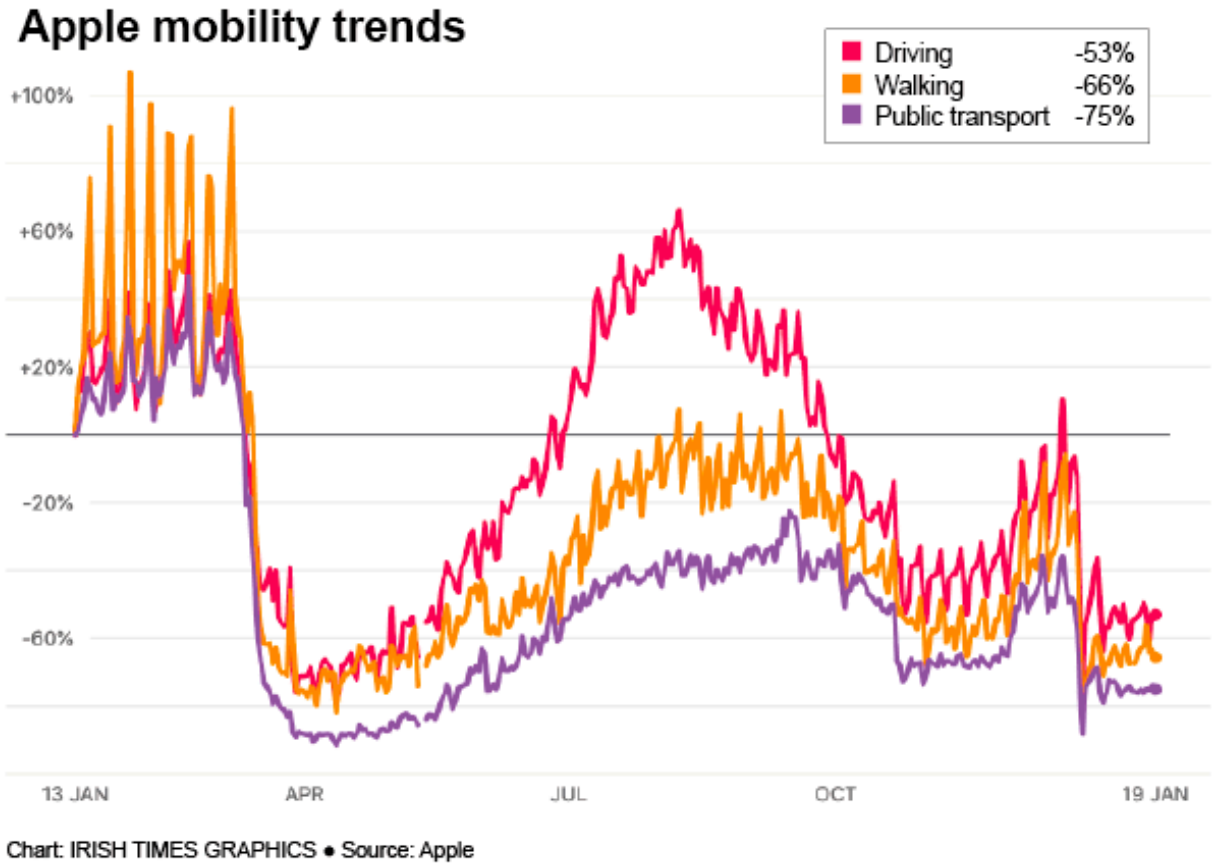
How does it compare with this?



https://commons.wikimedia.org/wiki/File:Change_in_Average_Temperature.svg#/media/File:Change_in_Average_Temperature.svg

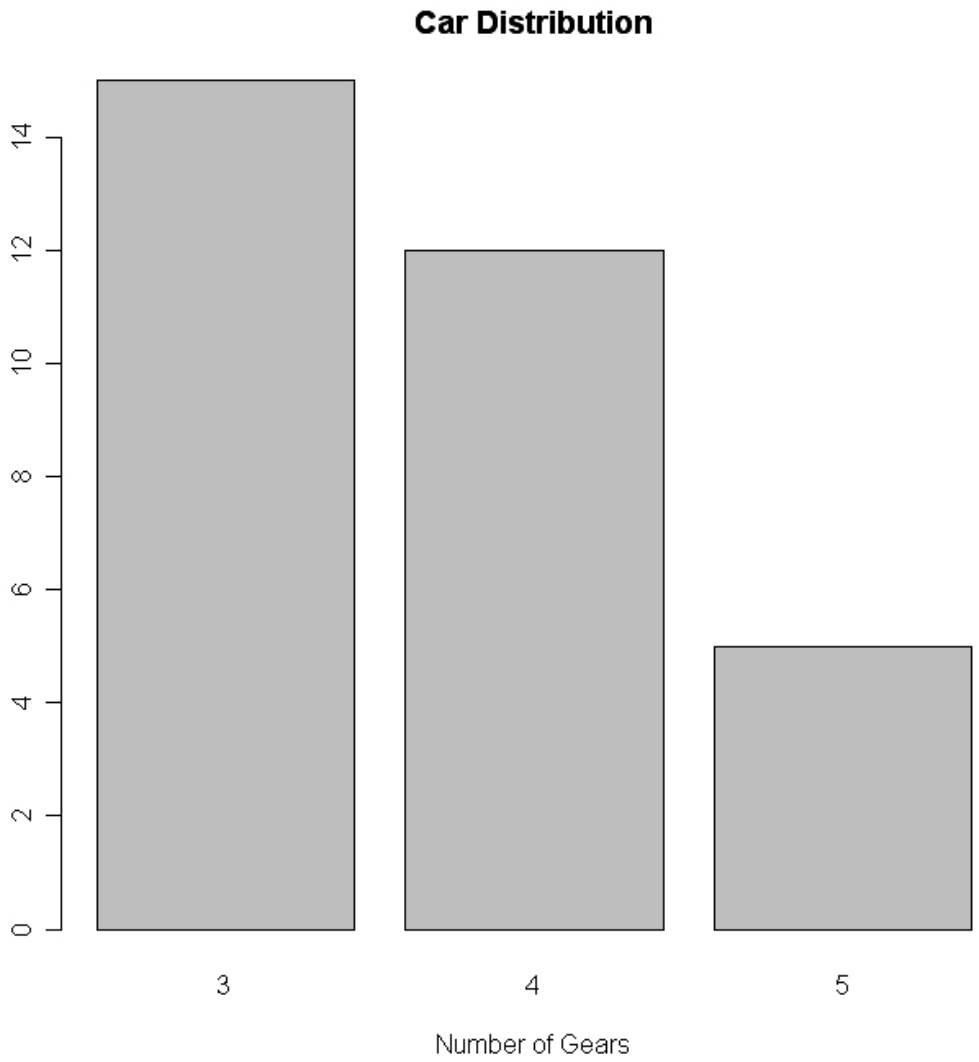
Trends over time

- Mobility patterns
- Dublin
- Jan 2020 to Jan 2021.

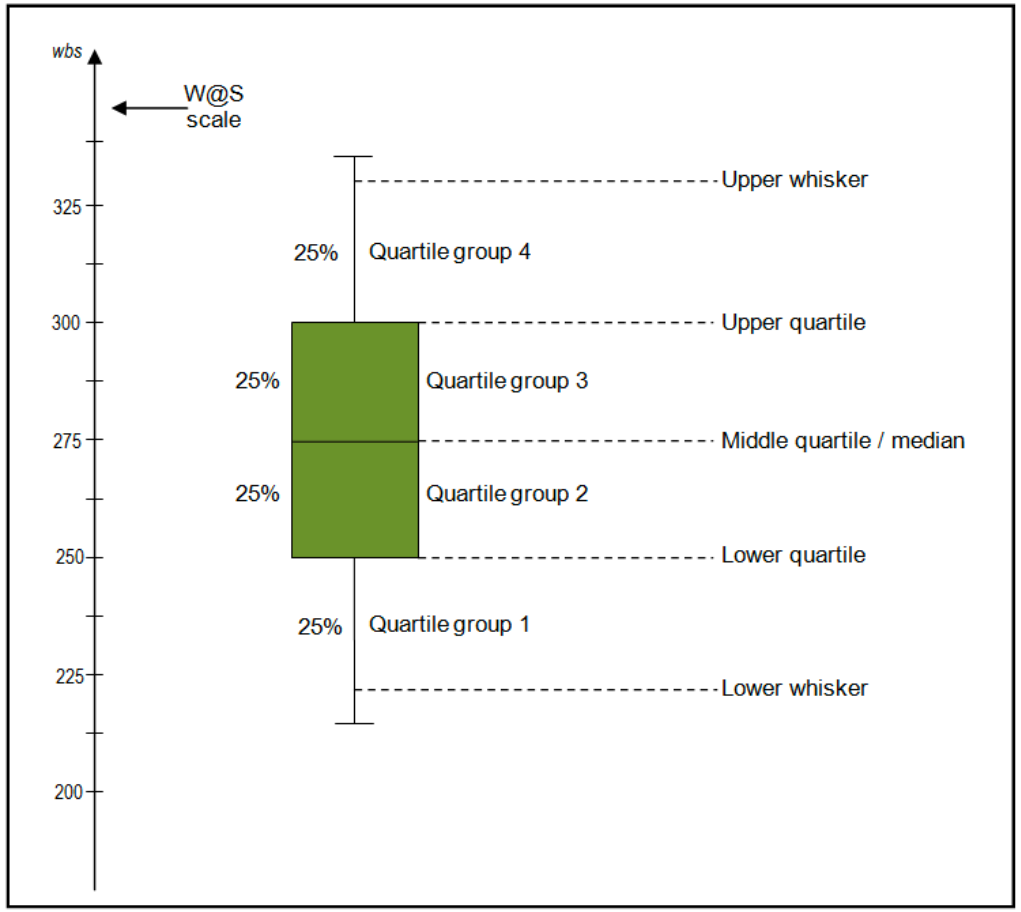


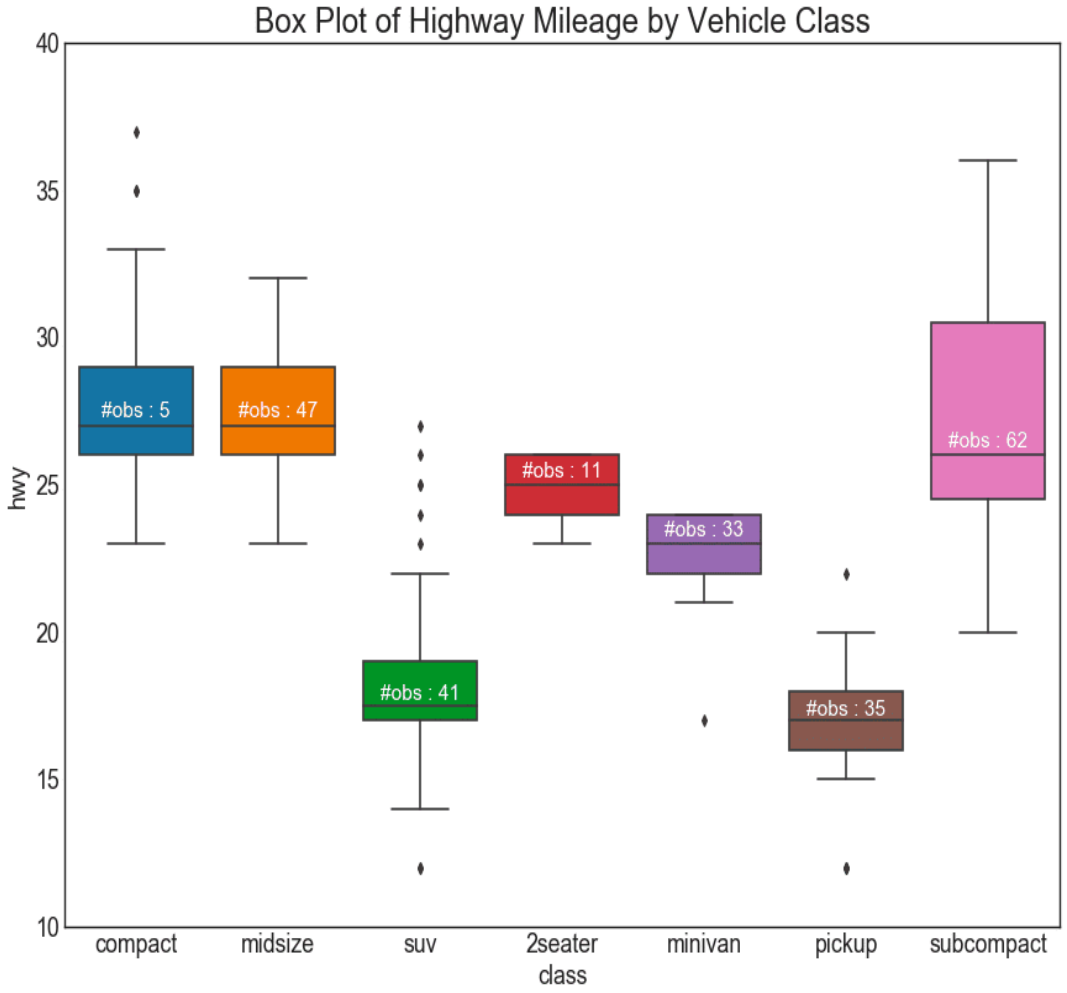
https://www.irishtimes.com/polopoly_fs/1.4464238!/image/image.png

Bar chart - Comparisons

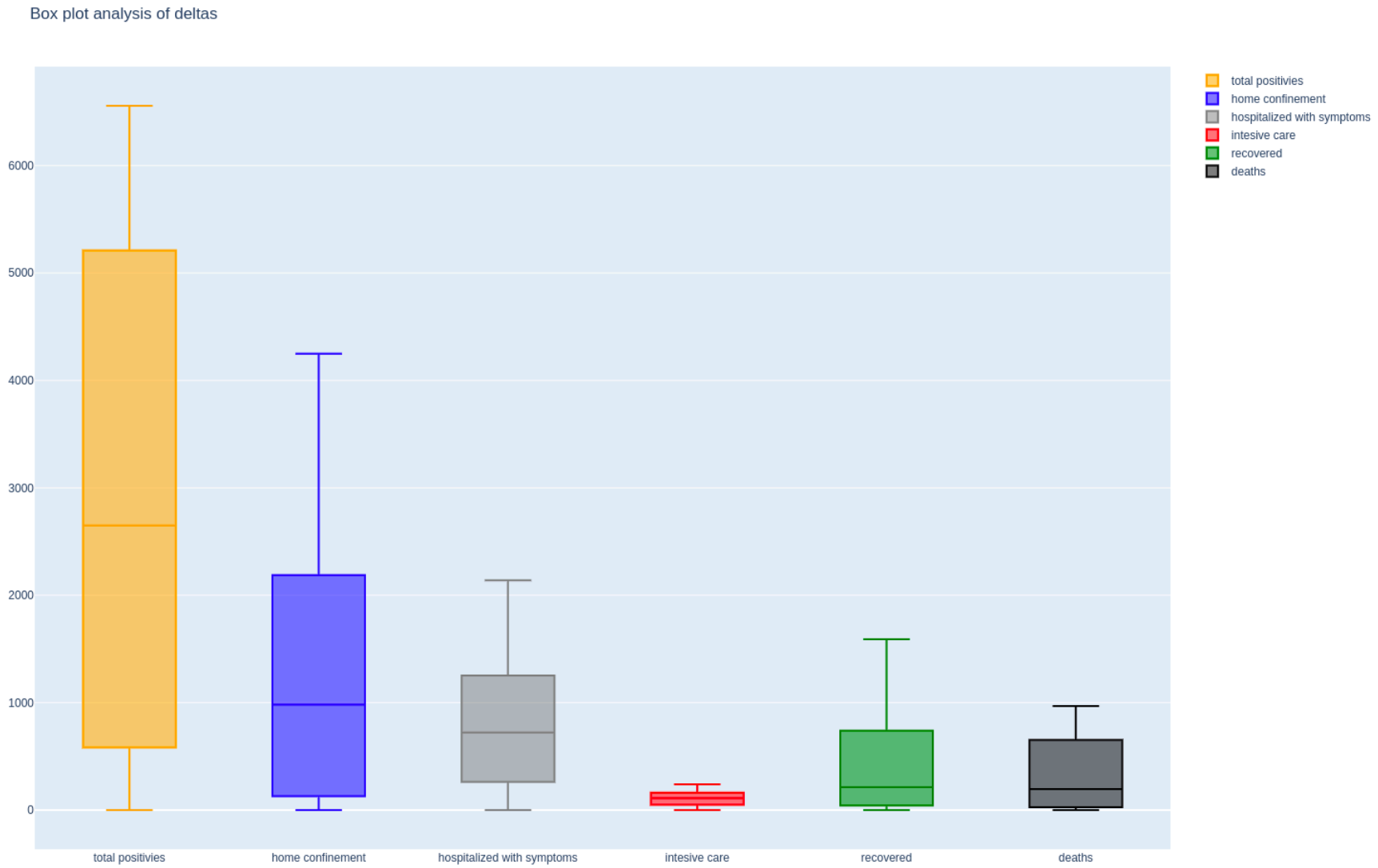


Box Plots – distribution of data





Covid cases in Italy

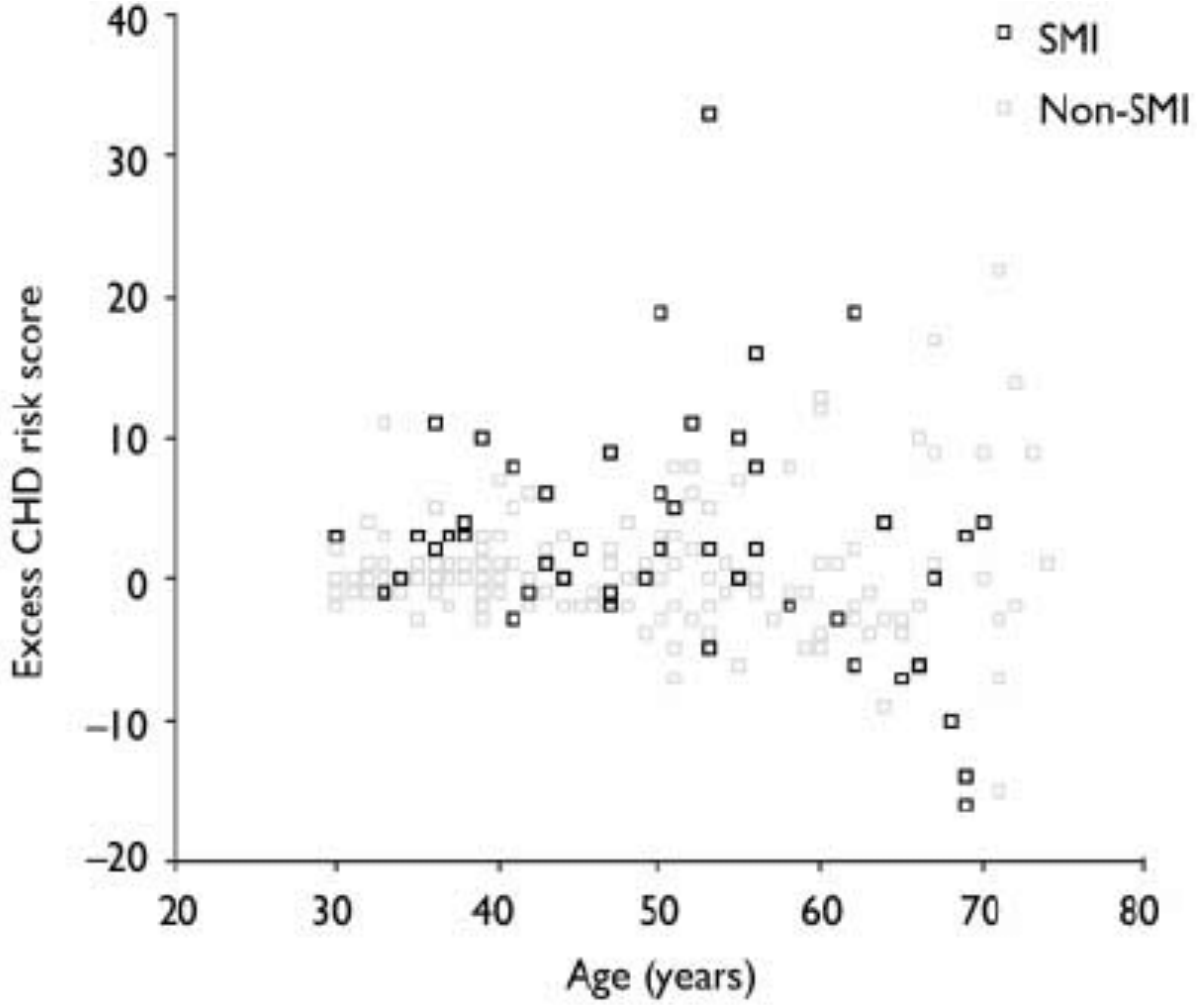


https://towardsdatascience.com/explorative-data-analysis-of-covid-19-data-in-italy-d5665cb62c5a

Relationships

Cardiac
Heart
Disease

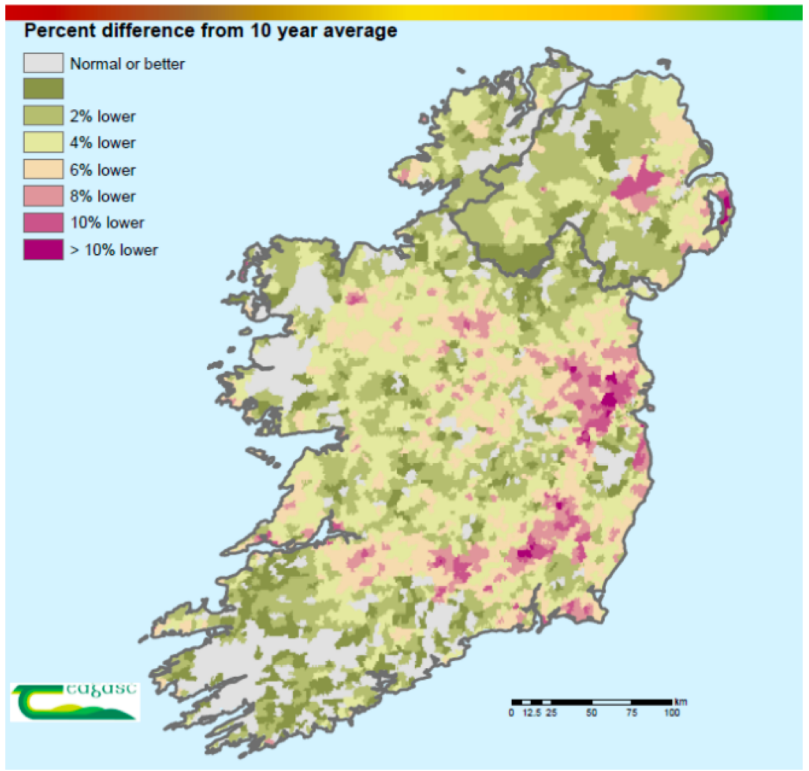
Serious
Mental
Illness



Maps

- Fodder produced in Ireland in 2018, compared to average over the years 2007 to 2017.

Poor fodder production in 2018 compared to average



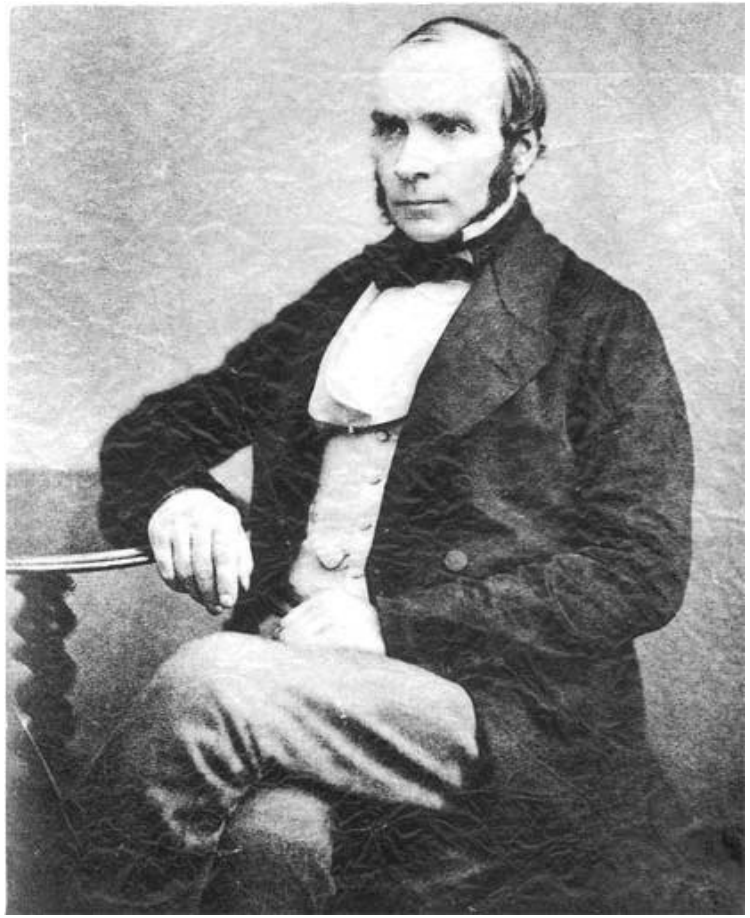
https://www.teagasc.ie/media/website/rural-economy/rural-economy/spatial-analysis/MOM_dec.pdf

Health InfoScape

http://senseable.mit.edu/healthinfoscape/

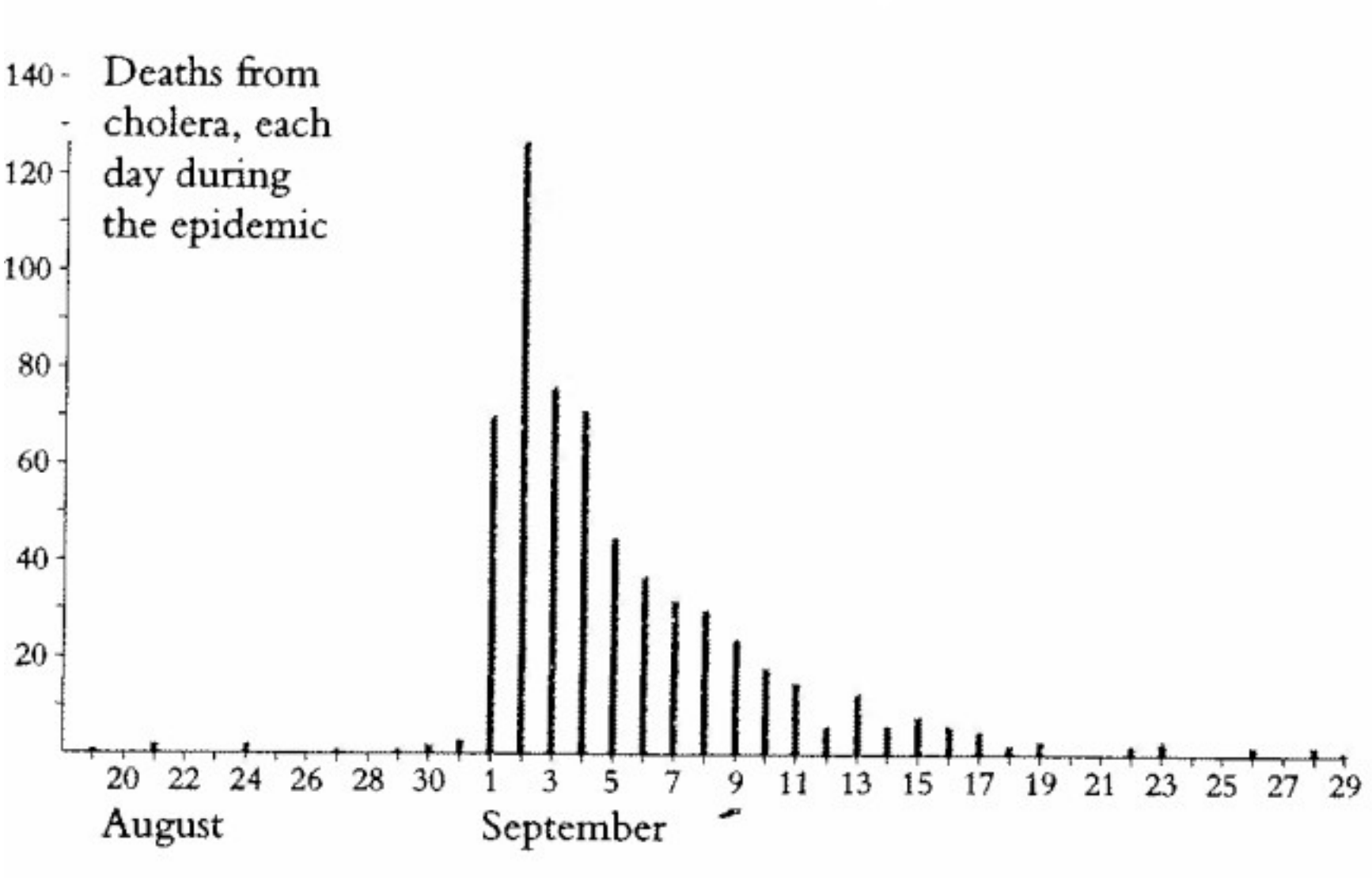
1854 London Cholera Epidemic

- On the evening of August 31, 1854 Cholera broke out in the Broad Street area of central London
- John Snow, a London doctor, was asked to investigate
- Visualisation was key to his solution

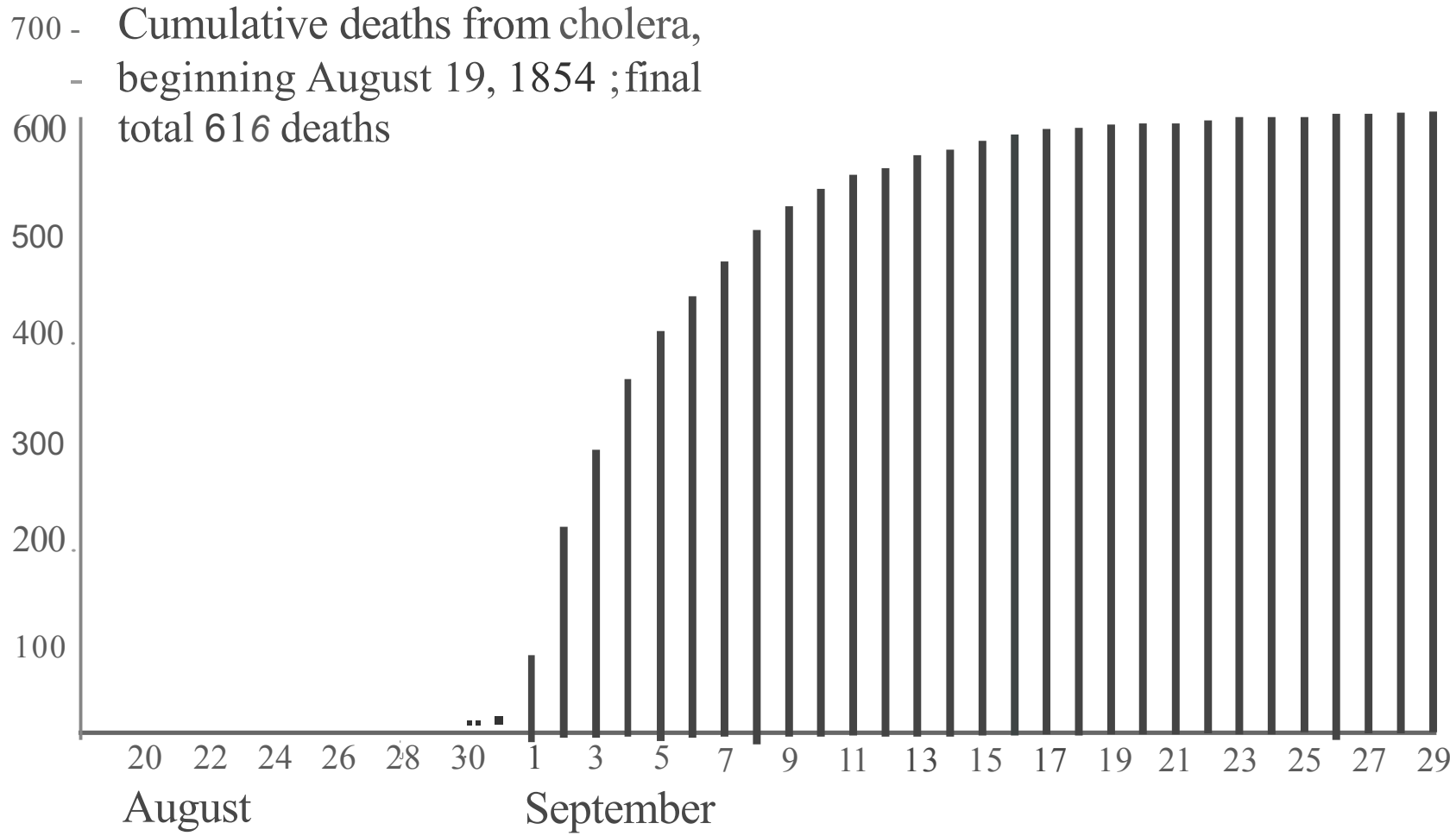


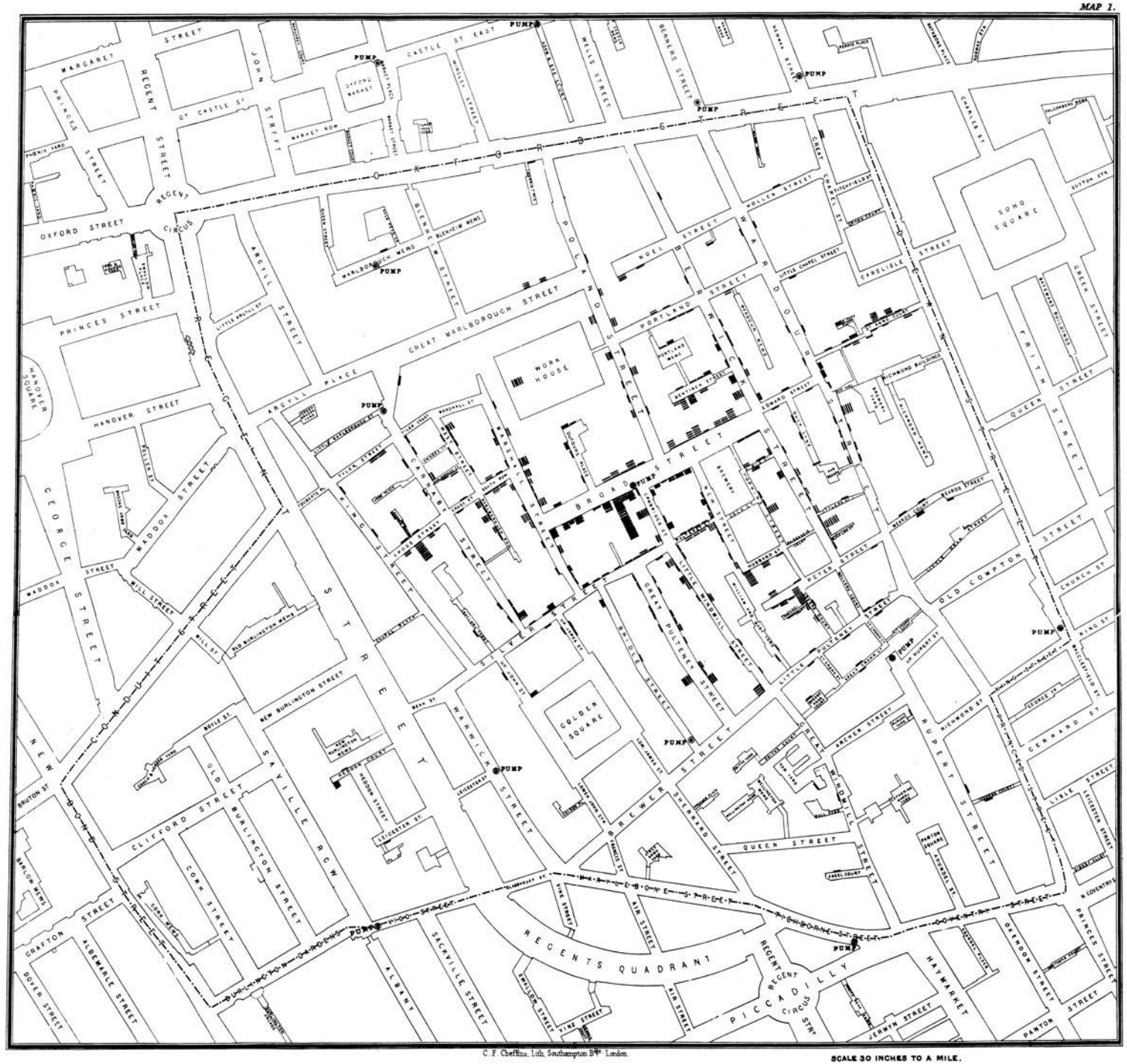
Adapted from “The Visual Display of Quantitative Information”, Graphics Press USA, 2001
www.edwardtufte.com/tufte/books_visex

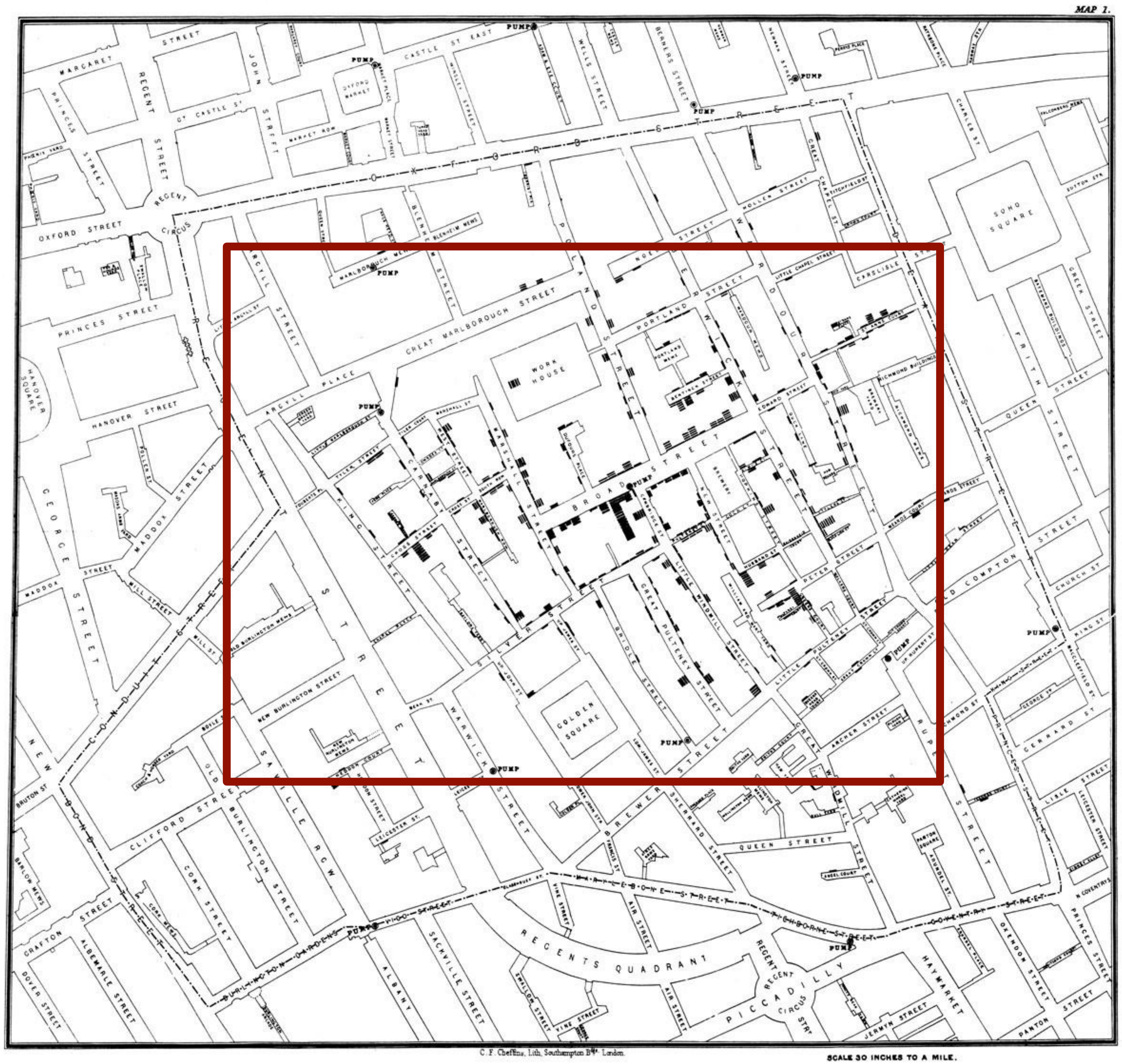
1854 London Cholera Epidemic

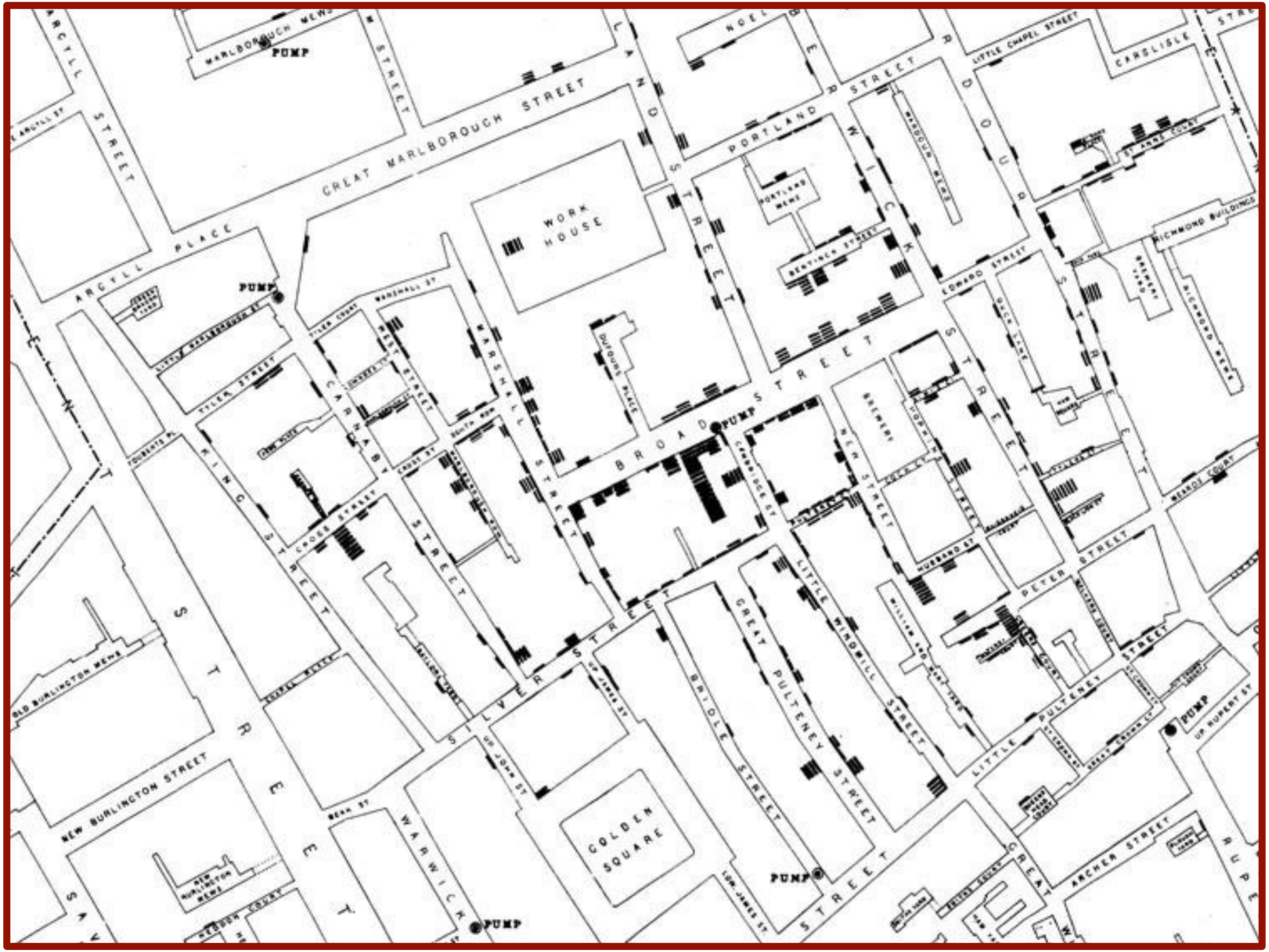


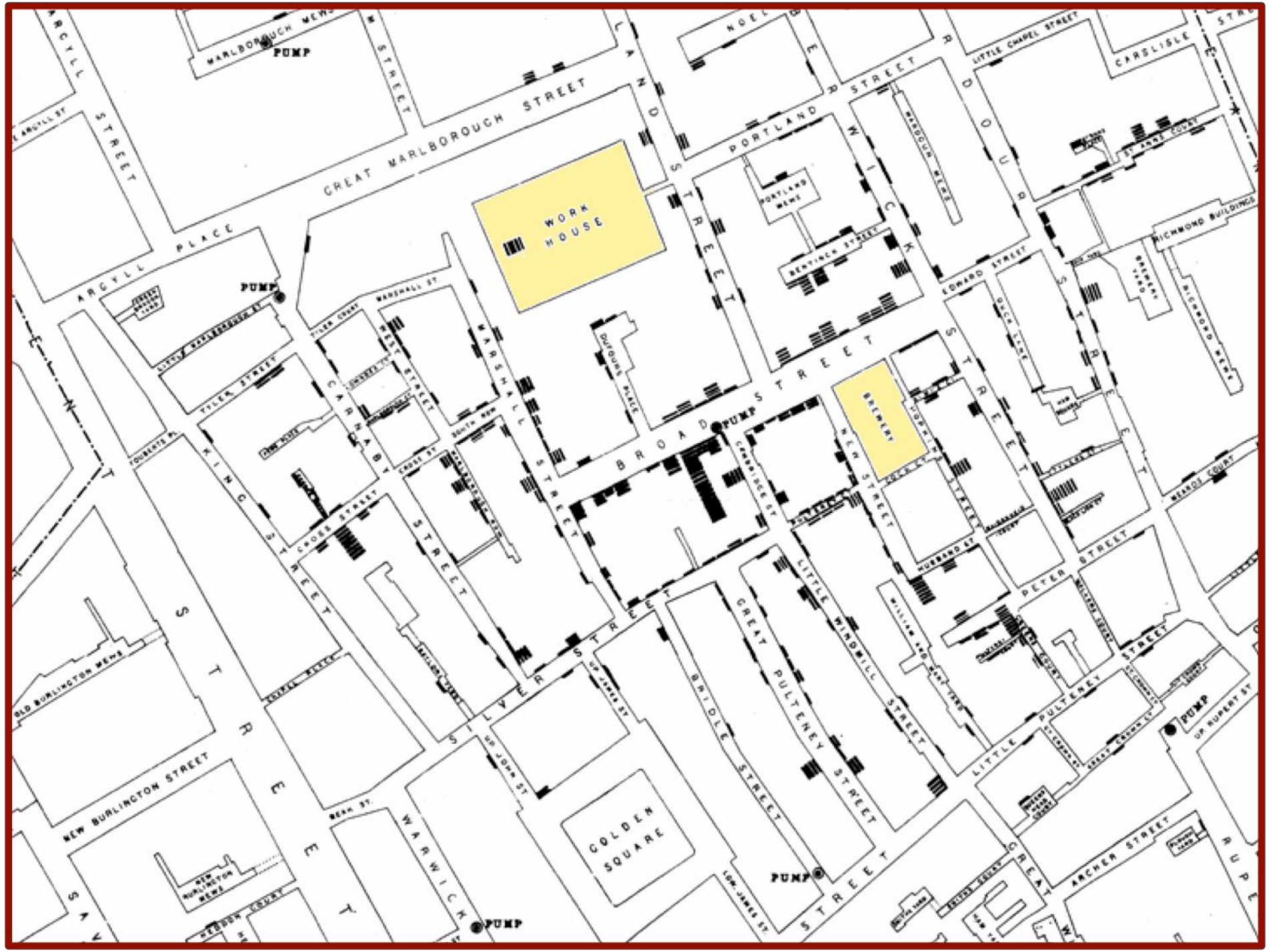
1854 London Cholera Epidemic

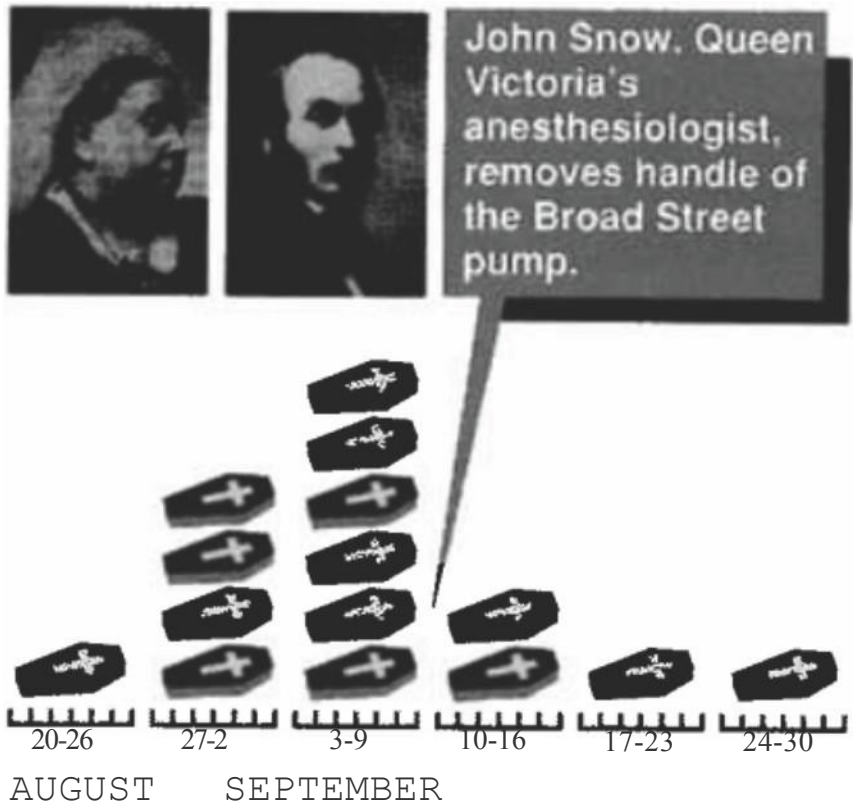



MAP 1.
PUMP
OXFORD STREET
REGENT STREET
CIRCUS
PRINCES STREET
HANOVER STREET
GEORGE STREET
MADDOX STREET
NEW BURLINGTON STREET
CLIFFORD STREET
SAVILLE ROW
BOYLE ST.
BRUTON ST.
CORK STREET
OLD BURLINGTON STREET
GRAFTON STREET
ALBEMARLE STREET
DOVER STREET
ARGYLL PLACE
ARGYLL STREET
GREAT MARLBOROUGH STREET
BLENHEIM STREET
MARLBOROUGH MEWS
POLAND STREET
WORK HOUSE
BROAD STREET
GOLDEN SQUARE
WARWICK STREET
LEICESTER ST.
BEAK ST.
GREAT PULTENEY STREET
BRIDLE STREET
LITTLE PULTENEY STREET
REGENTS QUADRANT
ALBANY
SACKVILLE STREET
REGENT CIRCUS
PICCADILLY
JERMYN STREET
HAYMARKET
PANTON STREET
PRINCES STREET
OXENDON STREET
RUPERT STREET
ARCHER STREET
QUEEN STREET
BREWER STREET
GERRARD ST.
LISLE STREET
LEICESTER STREET
OLD COMPTON STREET
QUEEN STREET
CHURCH ST.
WING ST.
GREEK STREET
FRITH STREET
SOHO SQUARE
CARLISLE STREET
SUTTON ST.
CHARLES ST.
FALCONBERG MEWS
GREAT WINDMILL STREET
BERNERS STREET
WELLS STREET
CASTLE ST EAST
GT CASTLE ST.
MARKET PLACE
MARKET ROW
OXFORD MARKET
JOHN STREET
REGENT STREET
MARGARET STREET
PRINCES STREET
PHOENIX YARD
HANOVER SQUARE
C. F. Cheffins, Lith. Southampton Bldgs. London.
SCALE 30 INCHES TO A MILE.

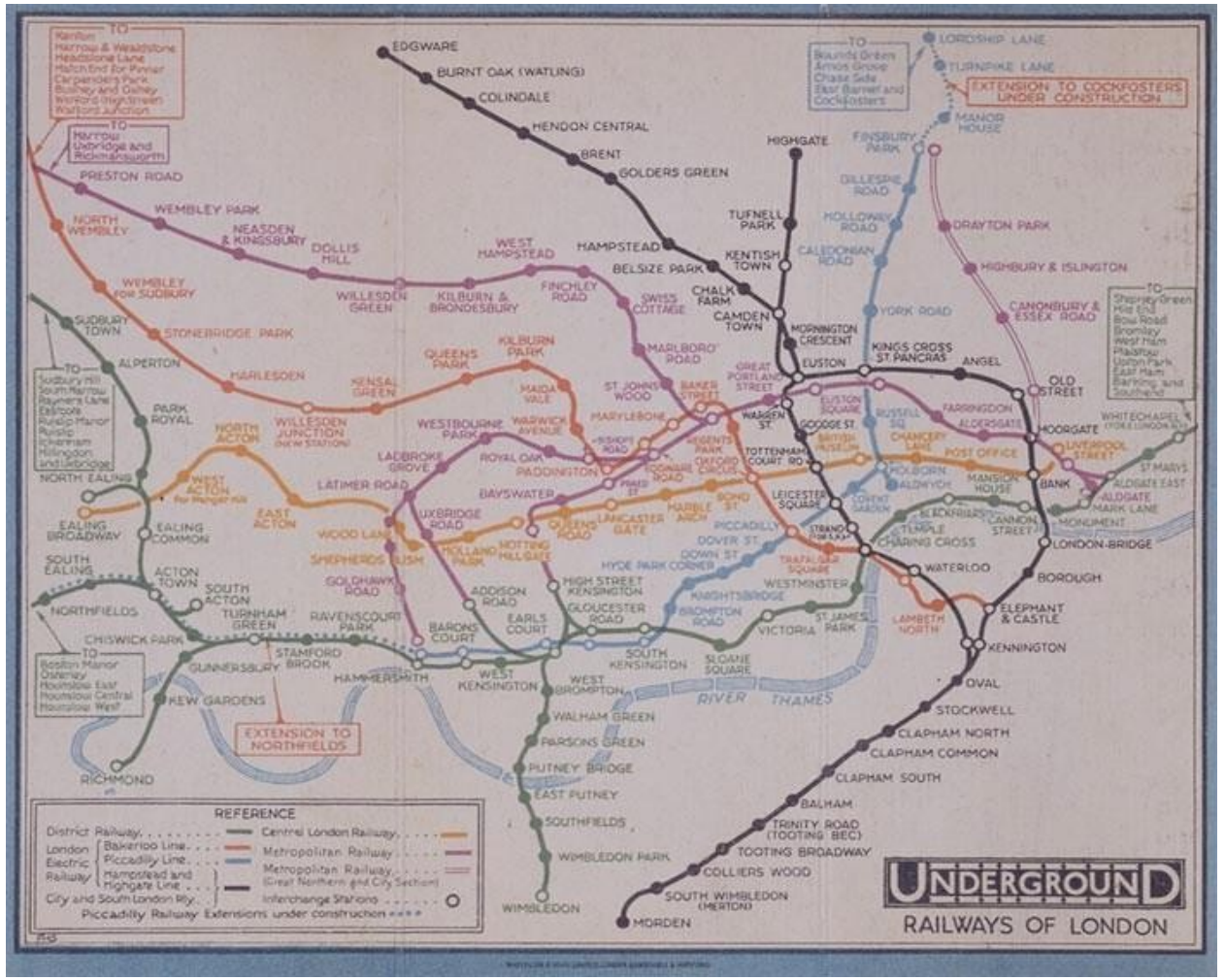

MAP 1.
OXFORD STREET
REGENT STREET
GREAT MARLBOROUGH STREET
BROAD STREET
WORK HOUSE
GOLDEN SQUARE
SOHO SQUARE
HANOVER SQUARE
REGENTS QUADRANT
PICCADILLY
QUEEN STREET
SAVILLE ROW
PUMP
C.F. Cheffins, Lith, Southampton Bdgs London.
SCALE 30 INCHES TO A MILE.


PUMP
MARLBOROUGH MEWS
GREAT MARLBOROUGH STREET
ARGYLL PLACE
WORK HOUSE
PORTLAND STREET
PORTLAND MEWS
BENTINCK STREET
LITTLE CHAPEL STREET
CARLISLE
ST ANNE'S COURT
RICHMOND BUILDINGS
RICHMOND MEWS
EDWARD STREET
DUFOURS PLACE
BROAD STREET
CAMBRIDGE ST
NEW STREET
BREWERY
TYLER STREET
LITTLE MARLBOROUGH ST
CARNABY
MARSHALL STREET
KING STREET
CROSS STREET
SILVER STREET
HUSBAND ST
PETER STREET
GREAT PULTENEY STREET
LITTLE WINDMILL STREET
BRIDLE STREET
LITTLE PULTENEY STREET
ARCHER STREET
GREAT CROWN CT
UP RUPERT ST
GOLDEN SQUARE
UP JAMES ST
UP JOHN ST
LOW JAMES ST
WARWICK STREET
BEAK ST
NEW BURLINGTON STREET
NEW BURLINGTON MEWS
OLD BURLINGTON MEWS
HEDDON COURT
MEARDS COURT


GREAT MARLBOROUGH STREET
ARGYLL PLACE
ARGYLL STREET
PUMP
WORK HOUSE
BROAD STREET
PORTLAND STREET
WARDOUR STREET
BREWERY
GOLDEN SQUARE
PUMP
PUMP
PUMP
PUMP
NEW BURLINGTON STREET
REGENT STREET
GREAT PULTENEY STREET
BRIDLE STREET
LITTLE PULTENEY STREET
PETER STREET
ARCHER STREET
WARWICK STREET
CARLISLE STREET
RICHMOND BUILDINGS
RICHMOND MEWS
PORTLAND MEWS
MARLBOROUGH MEWS
OLD BURLINGTON MEWS
NEW BURLINGTON MEWS


John Snow, Queen Victoria's anesthesiologist, removes handle of the Broad Street pump.
20-26
27-2
3-9
10-16
17-23
24-30
AUGUST
SEPTEMBER
=50 deaths
(Data rounded up to nearest whole coffin.)

London Underground Map



London Underground map circa 1930

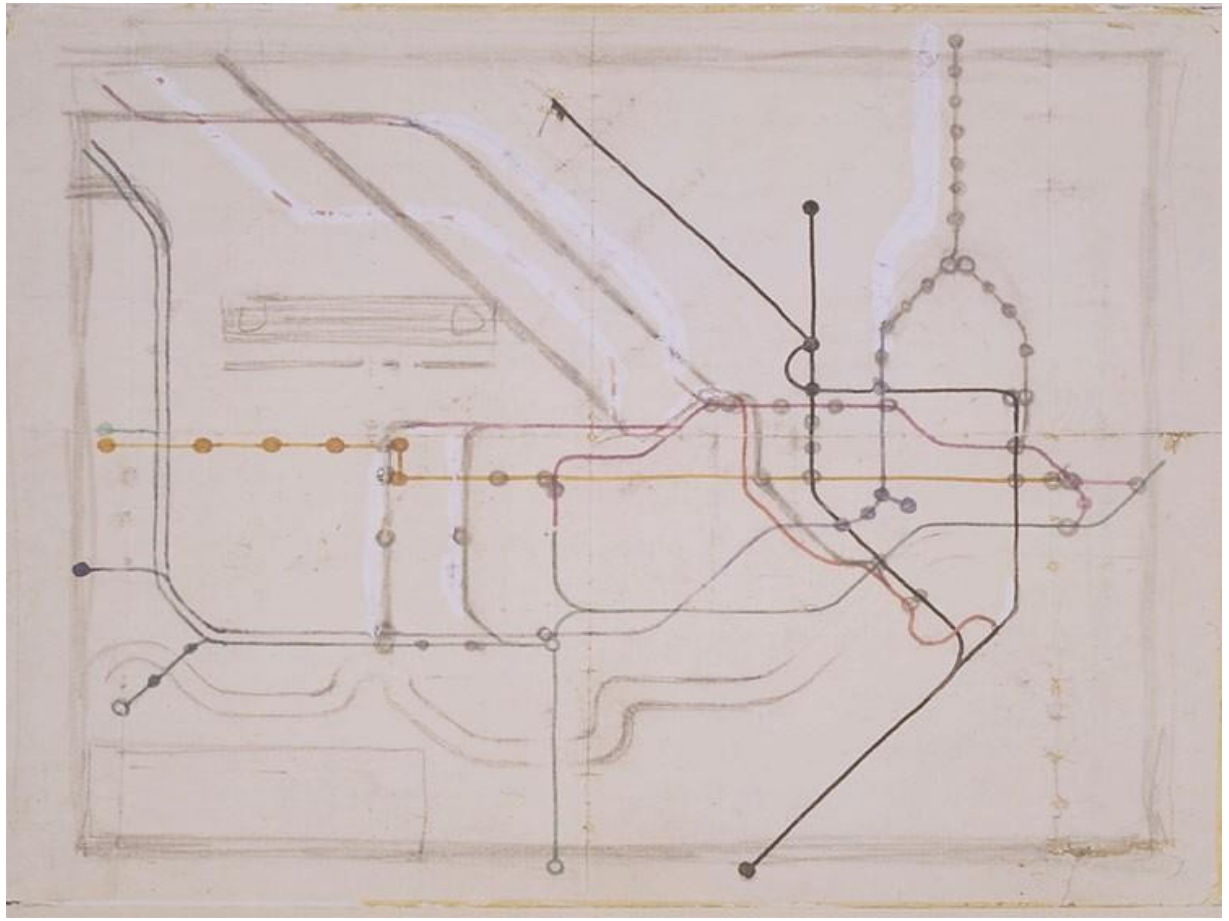
https://londonist.com/2016/05/the-history-of-the-tube-map

London Underground Map

Draughtsman Harry Beck, thought, *"If you're going underground, why do you need bother about geography? … Connections are the thing"* and designed a new map.

https://londonist.com/2016/05/the-history-of-the-tube-map

London Underground Map



An early sketch of Harry Beck's new map

https://londonist.com/2016/05/the-history-of-the-tube-map

London Underground Map

Beck's London Underground map circa 1933

https://londonist.com/2016/05/the-history-of-the-tube-map

Quiz

- Go to the POBYRNE room on Socrative
- After the video, answer the quiz questions.

Hans Rosling

https://youtu.be/jbkSRLYSojo

Visualisation?

- Well-designed visual representations can replace cognitive calculations with **simple perceptual inferences** and **improve comprehension**, **memory**, and **decision making**

- By making data more **accessible** and **appealing**, visual representations may also help engage more diverse audiences in exploration and analysis

- The challenge is to create **effective and engaging visualizations** that are **appropriate** to the data

"A Tour Through the Visualization Zoo", J. Heer, M. Bostock, & V. Ogievetsky, Communications of the ACM, vol. 53, no. 6, 2010

Visualisation?

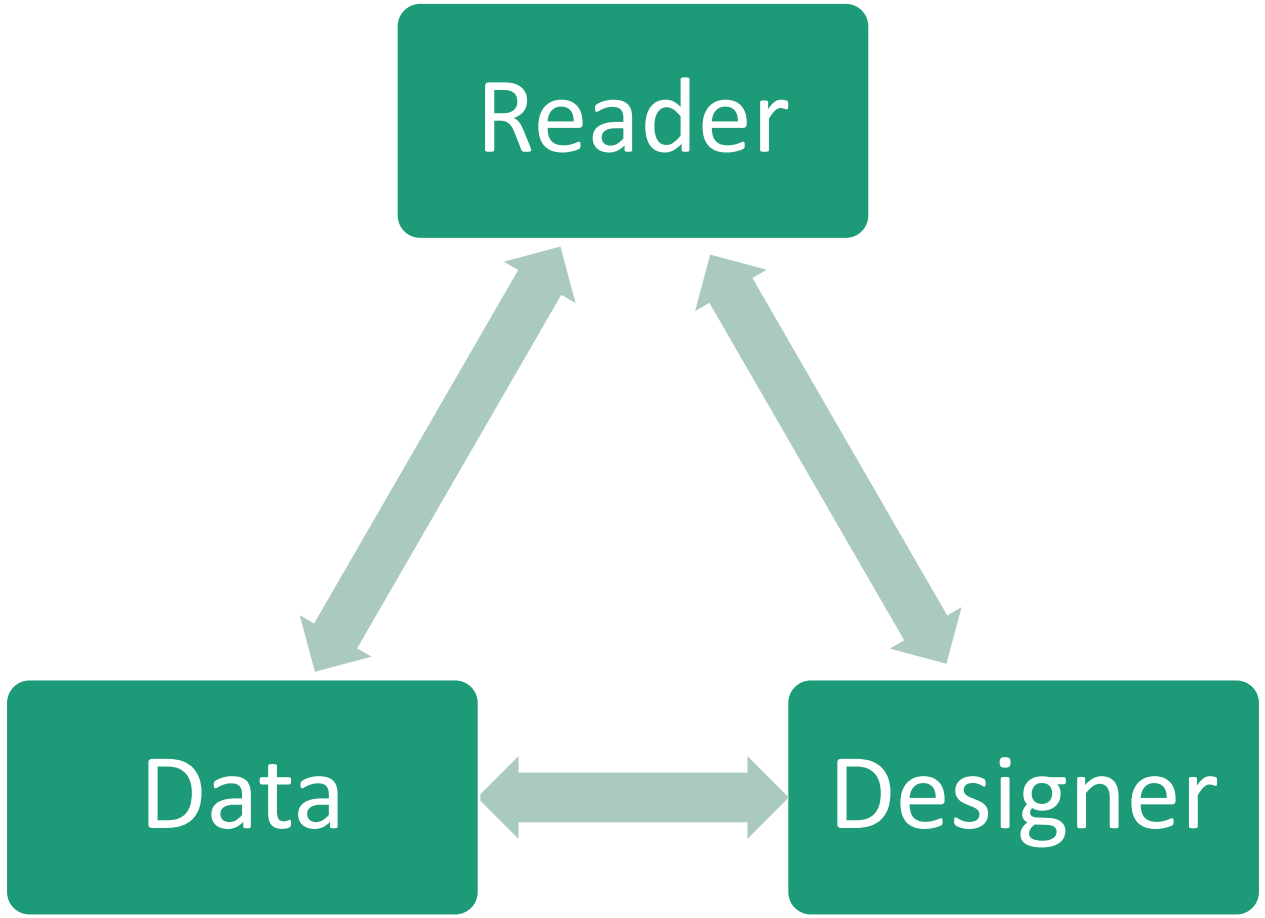
*The challenge is to create **effective** and **engaging** visualizations that are appropriate to the data.*

"A Tour Through the Visualization Zoo", J. Heer, M. Bostock, & V. Ogievetsky, Communications of the ACM, vol. 53, no. 6, 2010

Visualisation?

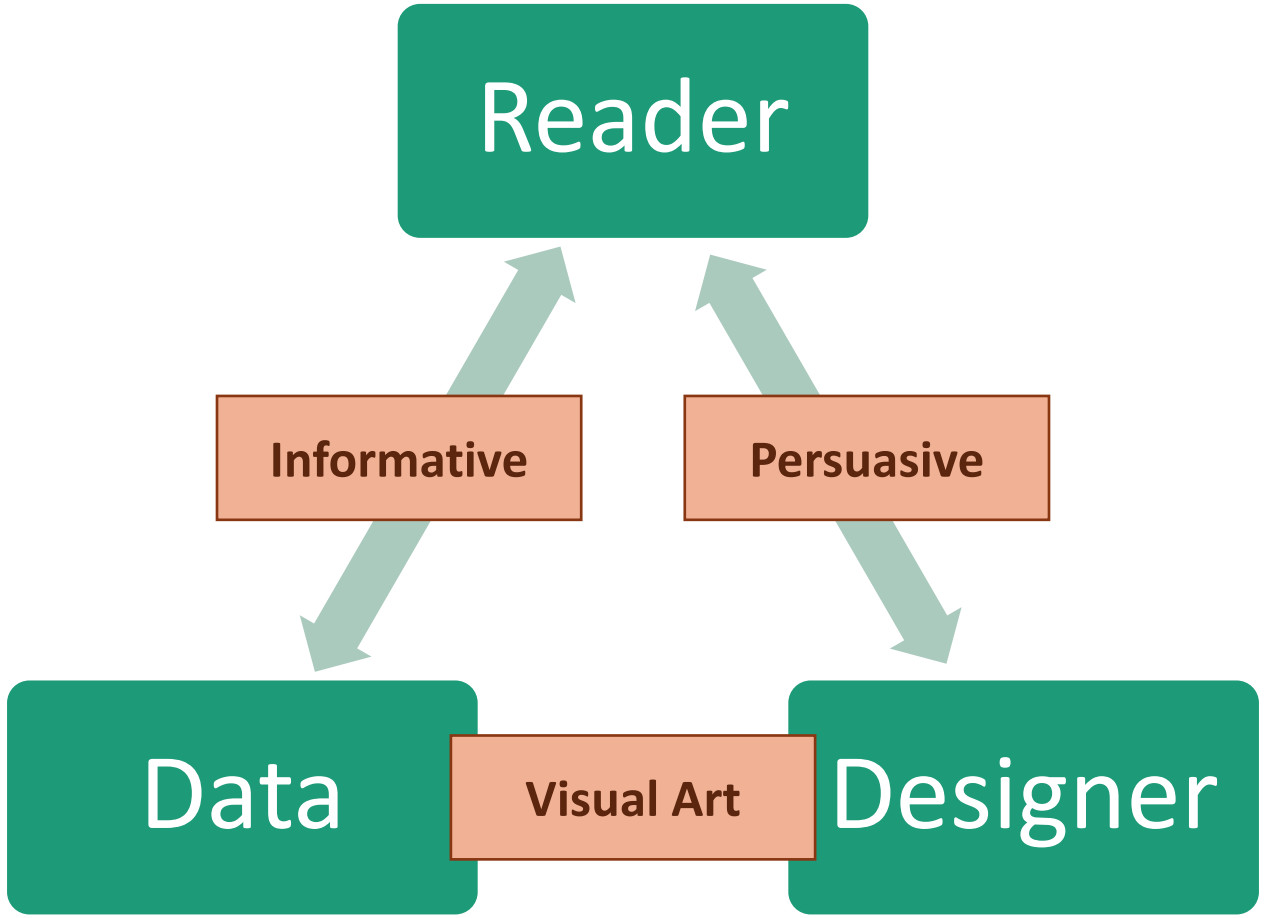
*The challenge is to create **effective** and **engaging** visualizations that are appropriate to the data,* **the audience and the message**

The Visualisation Trinity



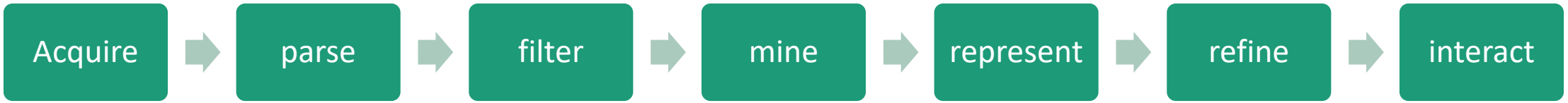
"Designing Data Visualizations", N. Iliinsky & J. Steele, O'Reilly Media, 2011

The Visualisation Trinity



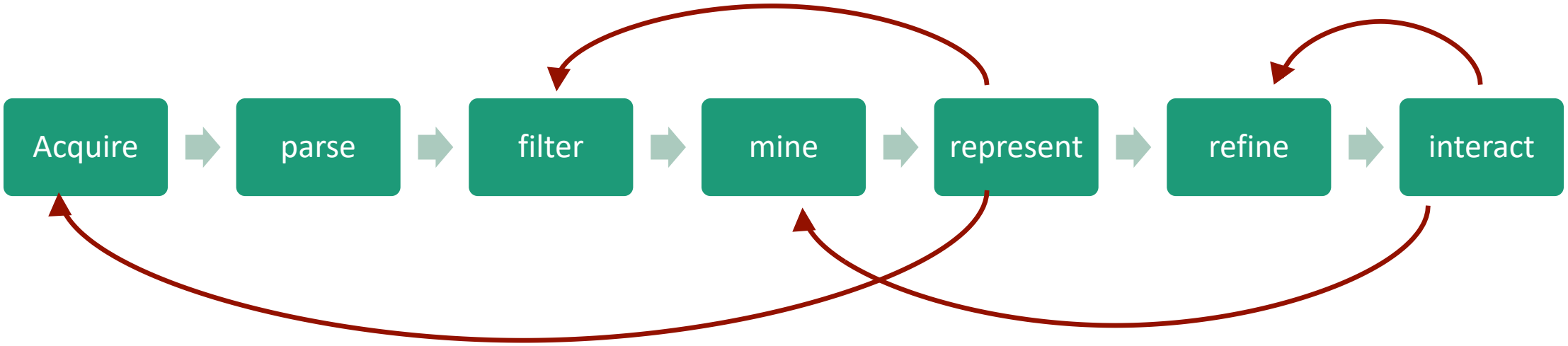
"Designing Data Visualizations", N. Iliinsky & J. Steele, O'Reilly Media, 2011

Ben Fry's Visualisation Steps



"Visualizing Data: Exploring and Explaining Data with the Processing Environment", B. Fry, O'Reilly Media, 2007

Ben Fry's Visualisation Steps



"Visualizing Data: Exploring and Explaining Data with the Processing Environment", B. Fry, O'Reilly Media, 2007

Module Overview

Module Aim

- The central aim of this module is help students attain the ability to take complex data, process it and extract value from it through visualisation

Learning Outcomes – as per Module Descriptor

- Demonstrate understanding of how humans perceive the world around them on a general level and absorb complex data/information on a specific level.
- Analyse and evaluate how metaphors are used to convey unfamiliar information.
- Analyse and evaluate how mental models aid in the interpretation of complex visual displays.
- Select, formulate and integrate metaphors to suit data-driven tasks
- Design effective visualizations based on principles from perceptual psychology, cognitive science, graphic design and visual art.
- Create and deploy successful data visualisations using leading software tools
- Demonstrate an understanding how visualisation is used in date journalism to communicate complex ideas and stories
- Demonstrate understanding how visualisation is used in story telling

Learning Outcomes

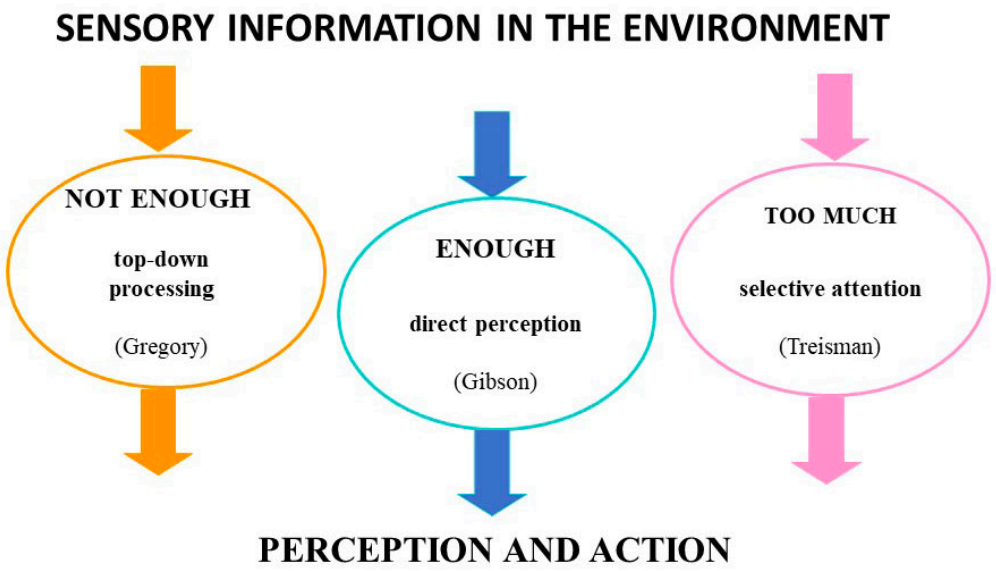
1. Critically analyse visualisations in terms of how humans perceive the world around them on a general level and absorb complex data/information on a specific level
2. Apply and interpret how metaphors are used to convey unfamiliar information
3. Select, formulate and integrate metaphors to suit data-driven tasks
4. Design effective visualizations based on principles from perceptual psychology, cognitive science, graphic design and visual art.
5. Create and deploy successful data visualisations using leading software tools
6. Create, deploy and critically analyse visualisations as used in story telling

What do we see?

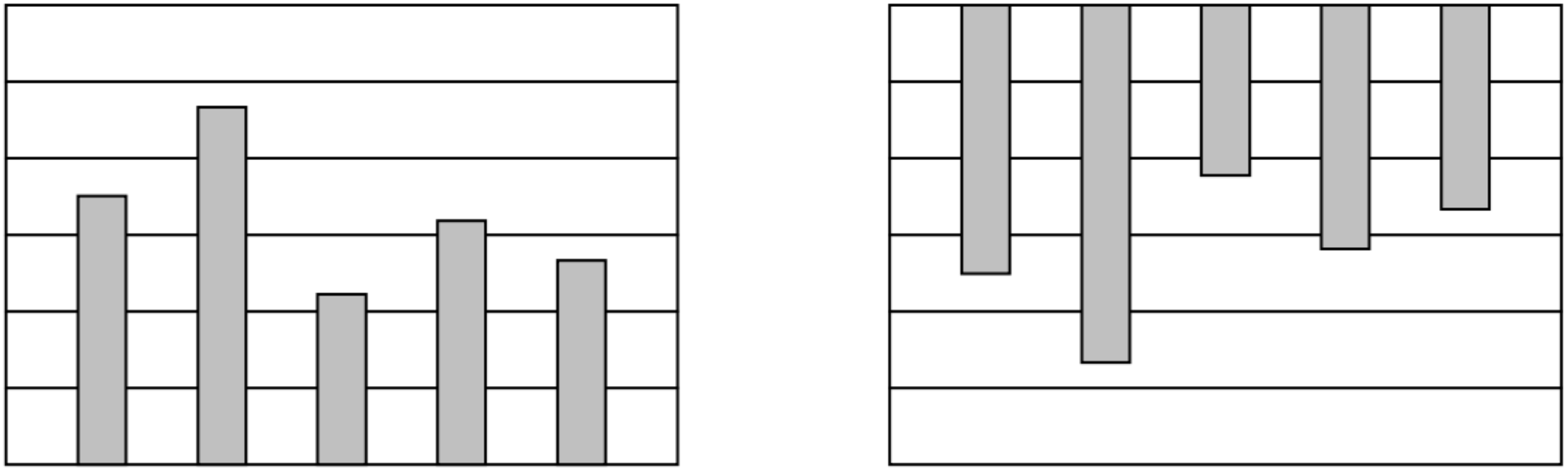
1. Critically analyse visualisations in terms of how humans
   - perceive the world around them on a general level

What do we see?

1. Critically analyse visualisations in terms of how humans
   - perceive the world around them on a general level
   - and absorb complex data/information on a specific level



2. How metaphors convey information

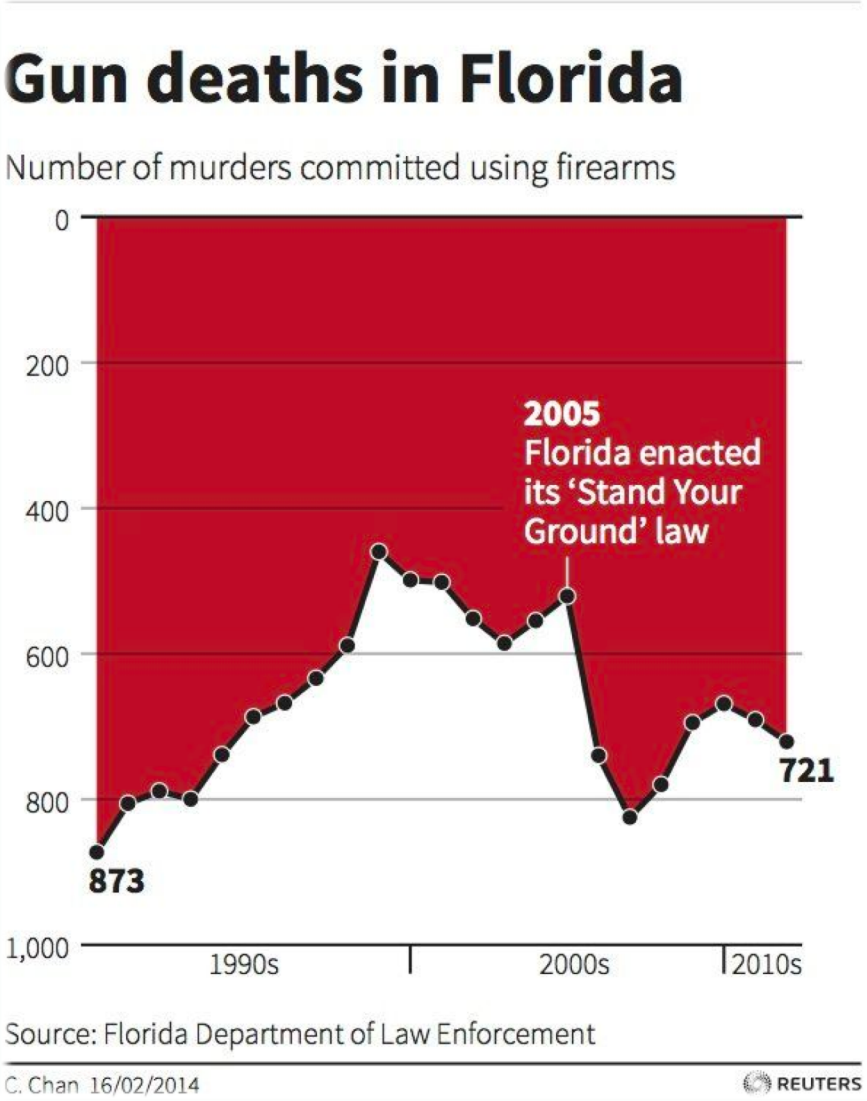


(a) (b)

Figure 1. Comparison of otherwise identical bar charts employing MORE IS UP (a) vs. MORE IS DOWN (b) metaphorical mappings. Subjectively, (a) is quantitatively and comparatively easier to read, and feels more "intuitively correct," than (b).

How metaphors convey information

- Did the law decrease deaths?



3. Picking metaphors

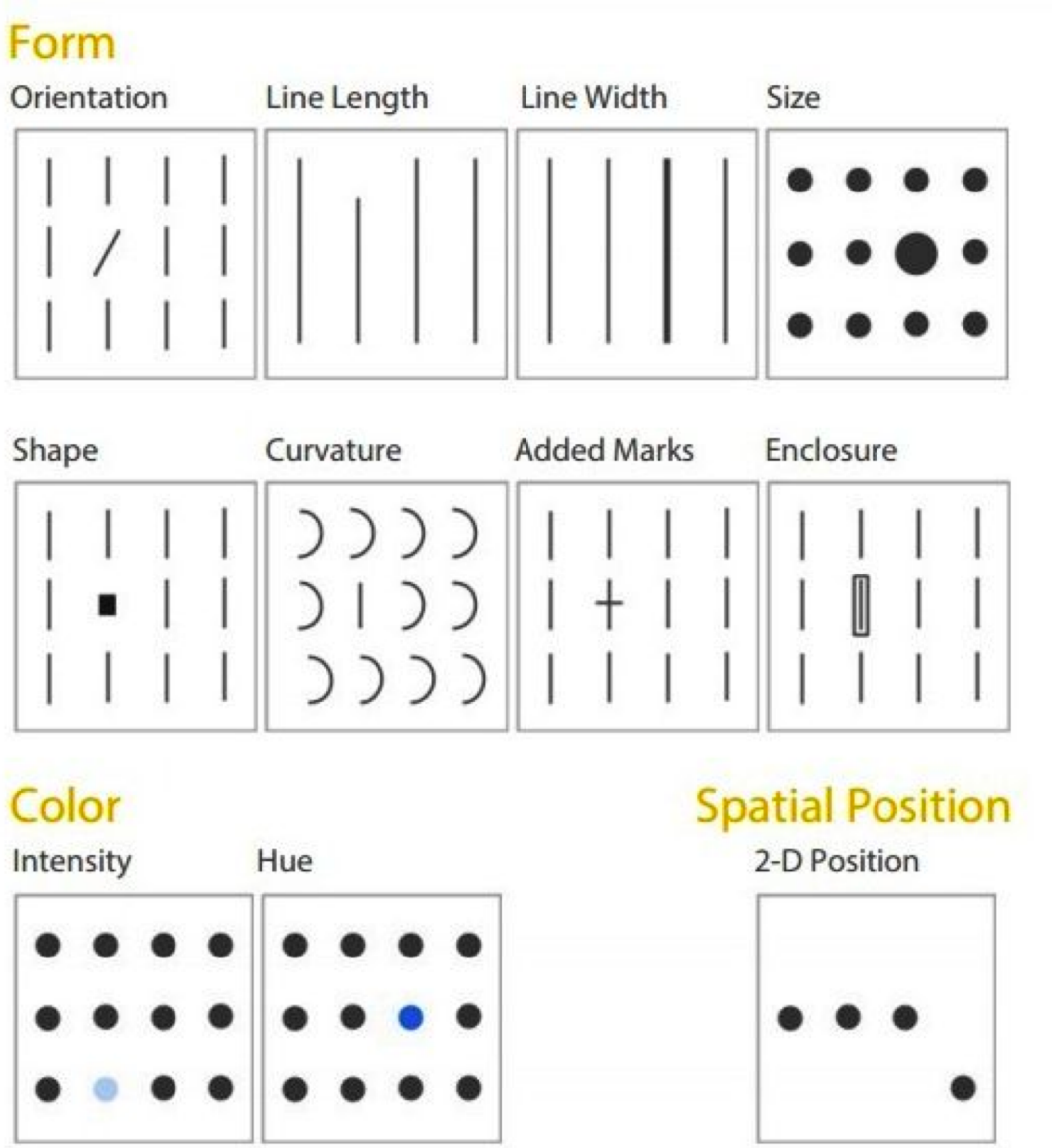
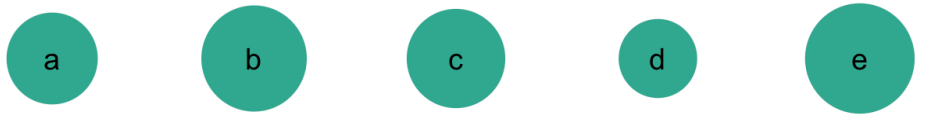
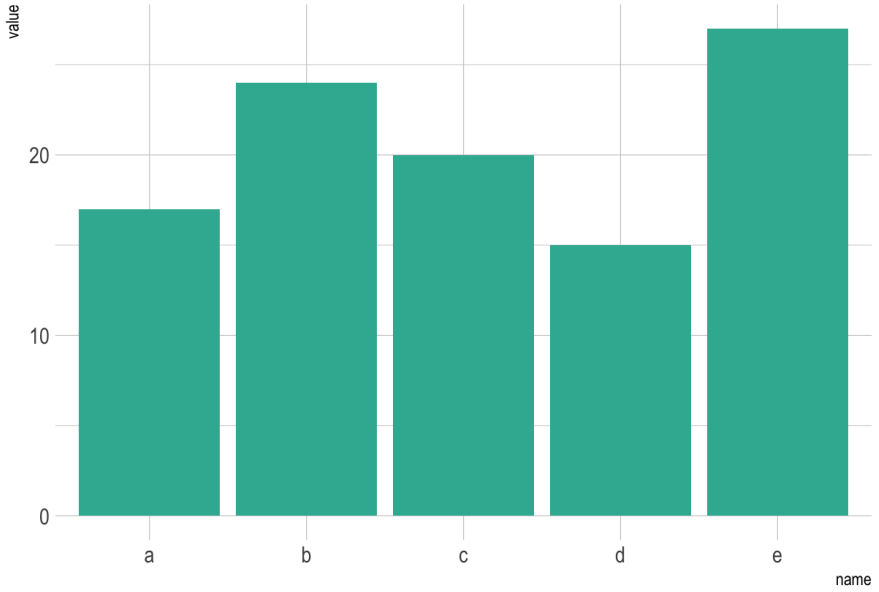


Image credit: Colin Ware, [Information Visualization, Third Edition: Perception for Design]

3. Picking metaphors

- Which visualisation compares area better?





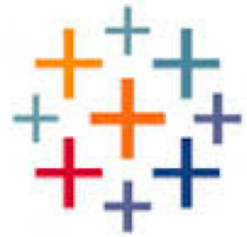
https://www.data-to-viz.com/caveat/area_hard.html

4. Design effective visualizations

- Using
  - perceptual psychology,
  - cognitive science,
  - graphic design and
  - visual art.

5. Create successful data visualisations

- Using leading software tools.



Tools

+ableau+++public

- Tableau is an enterprise level visualisation-based business analytics tool
- Tableau Public is its little brother!
- www.tableausoftware.com/products/ public

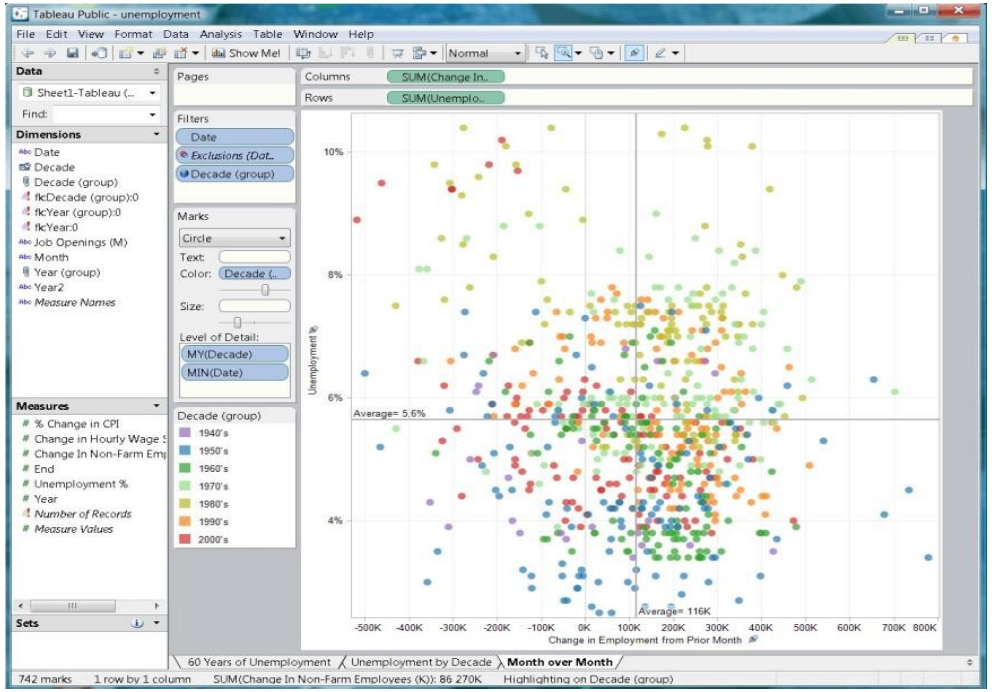
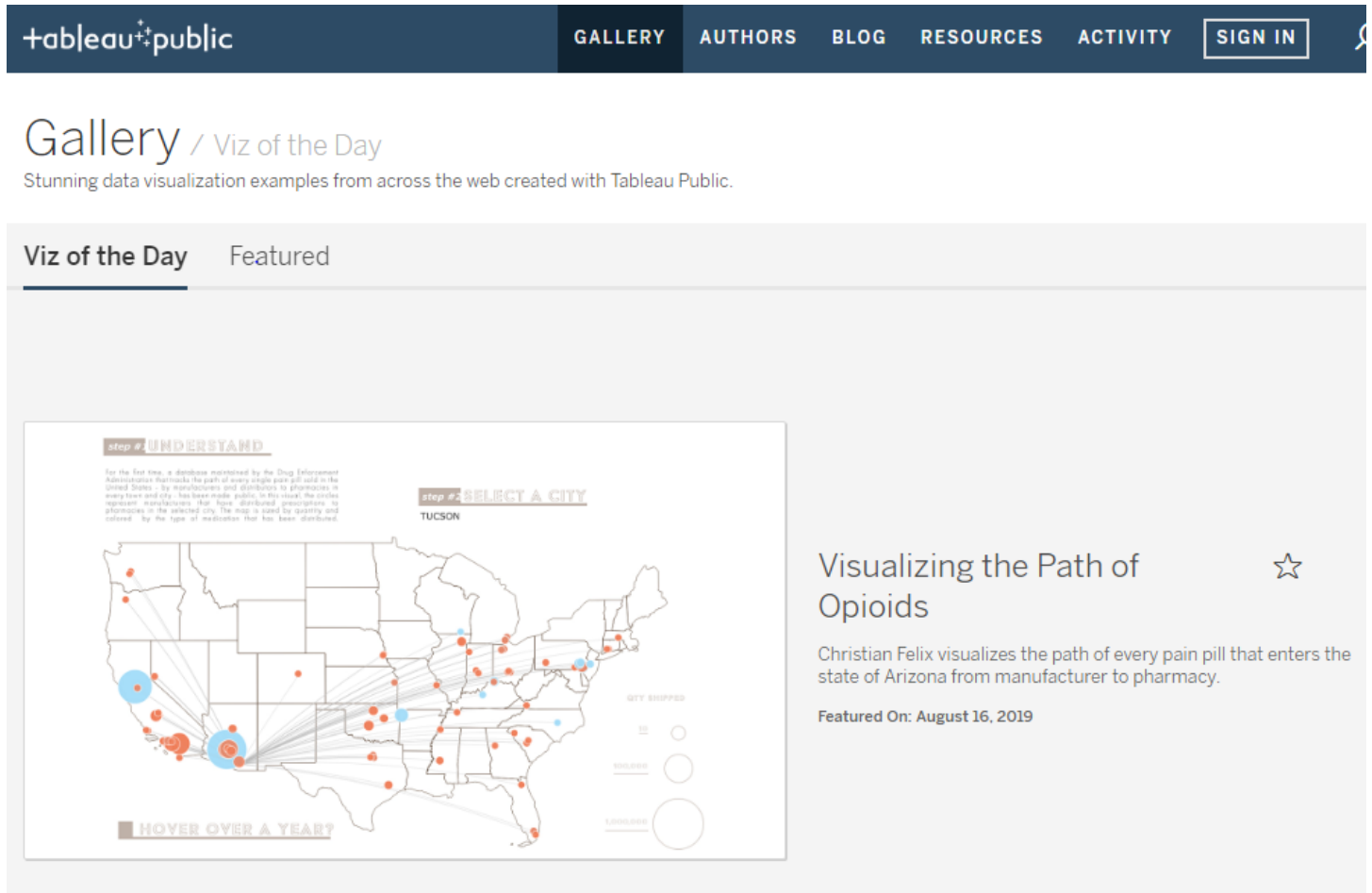


Tableau Public Visualisations

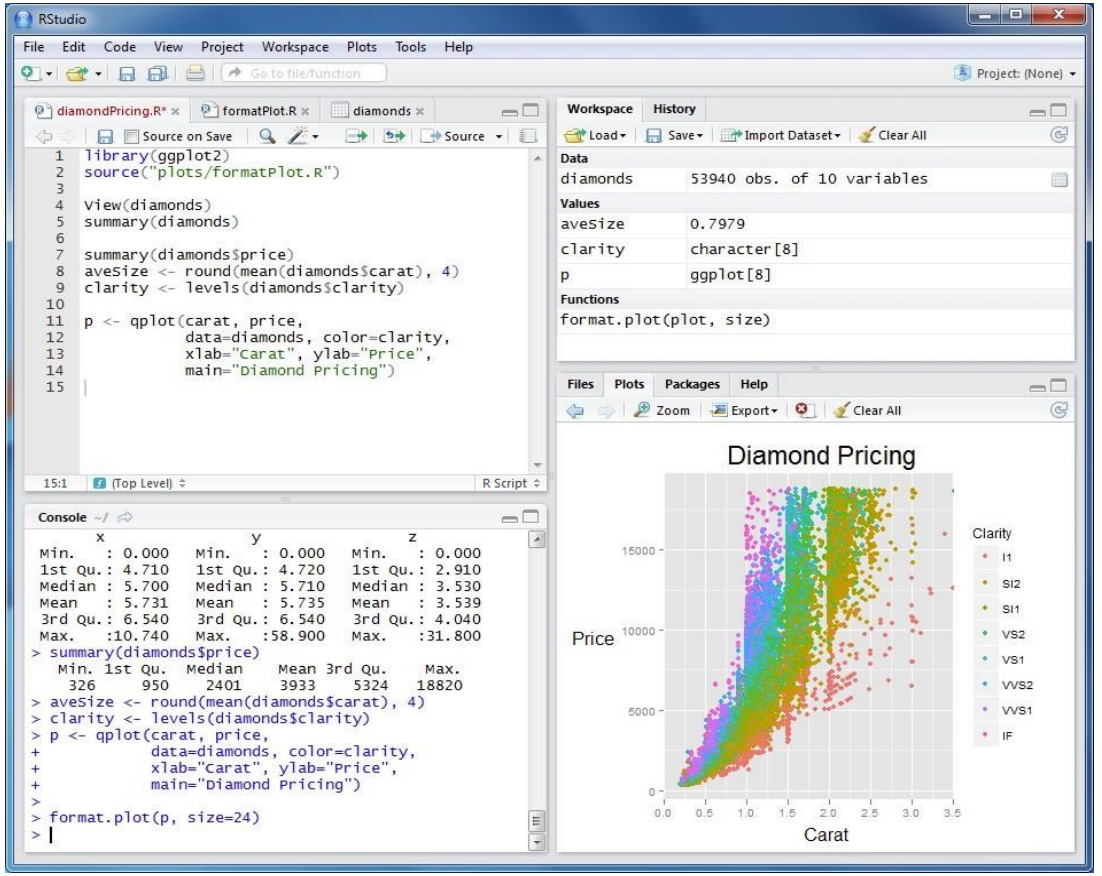


A gallery of visualisations created using Tableau Public

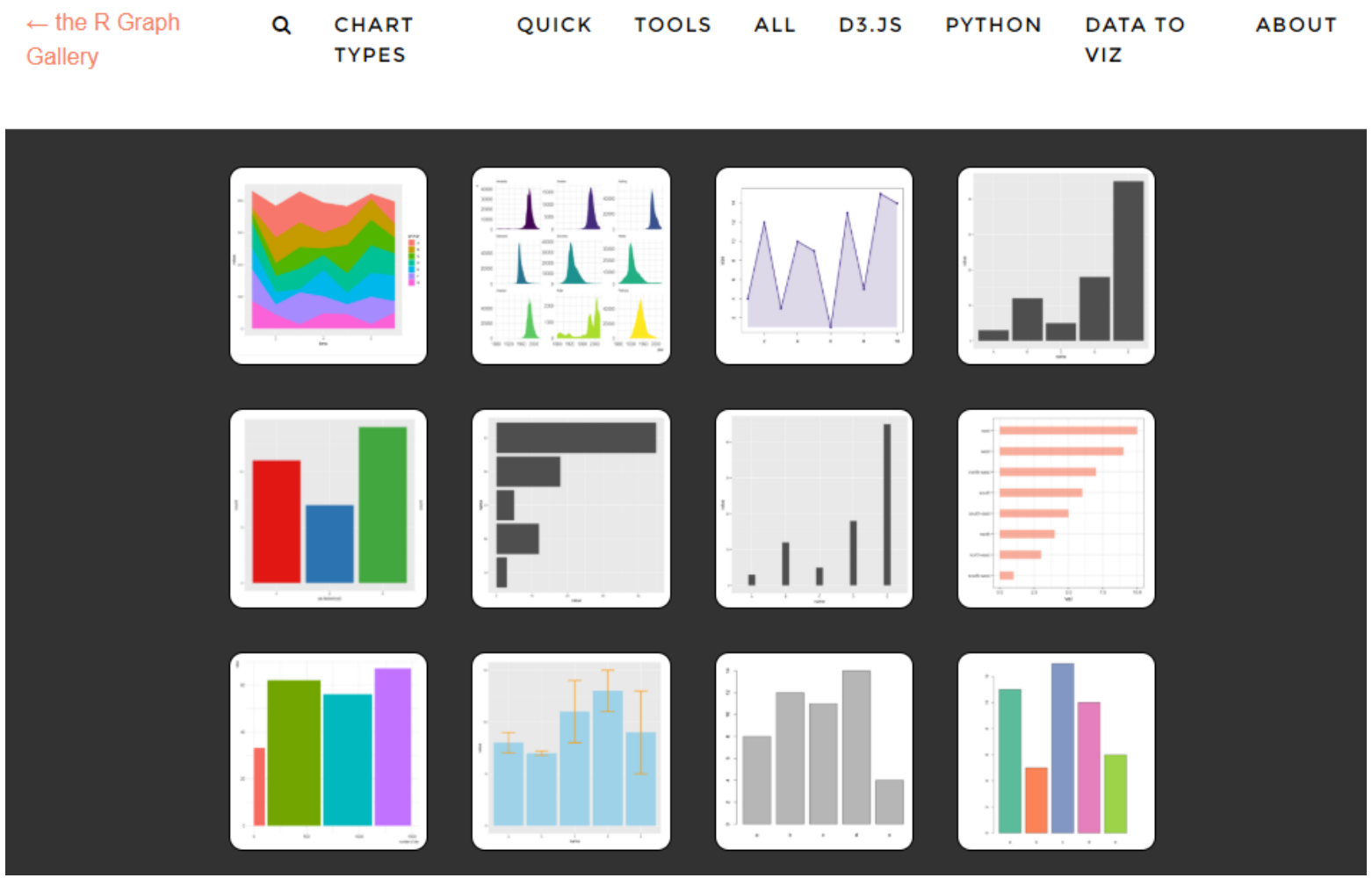
https://public.tableau.com/en-us/gallery

Tools

- R is a statistical programming
- language with extensive visualisation features
- R Studio is a rich GUI interface to R
- www.r-project.org
- www.rstudio.org



R Visualisations

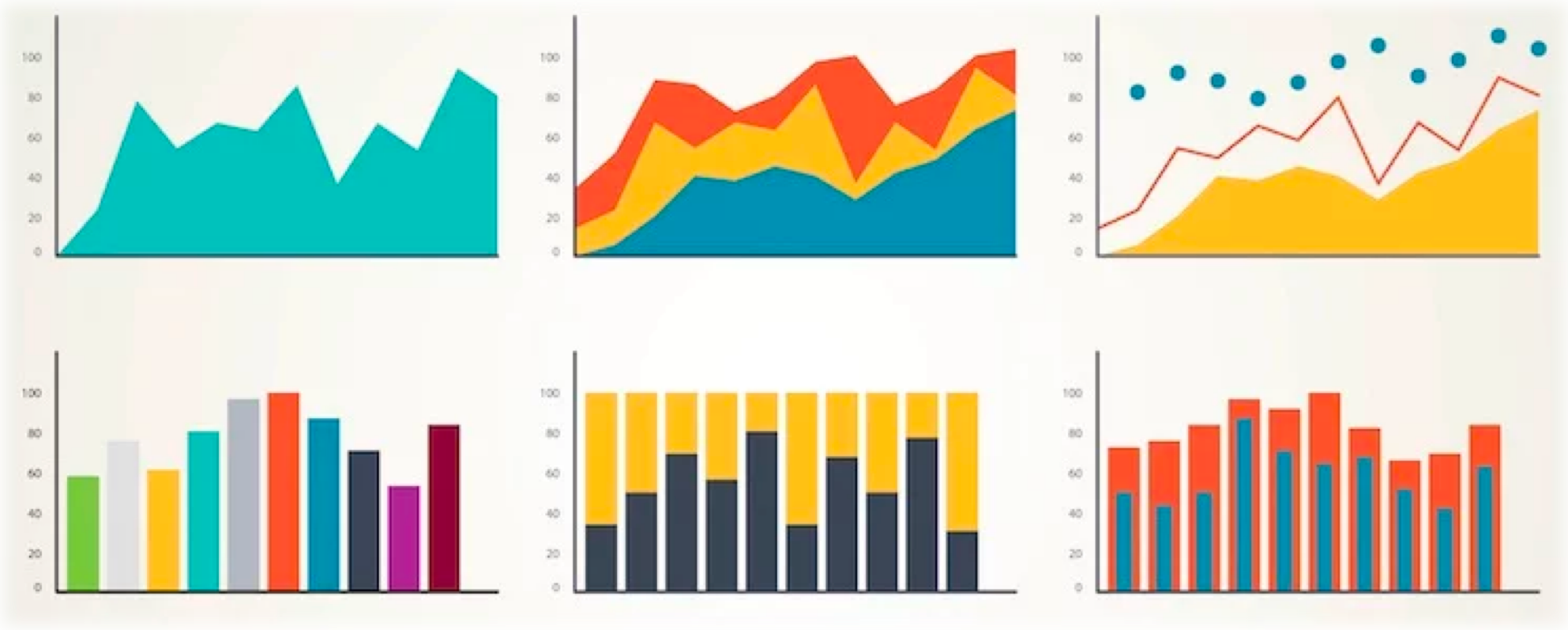


A gallery of visualisations created using R with source code!

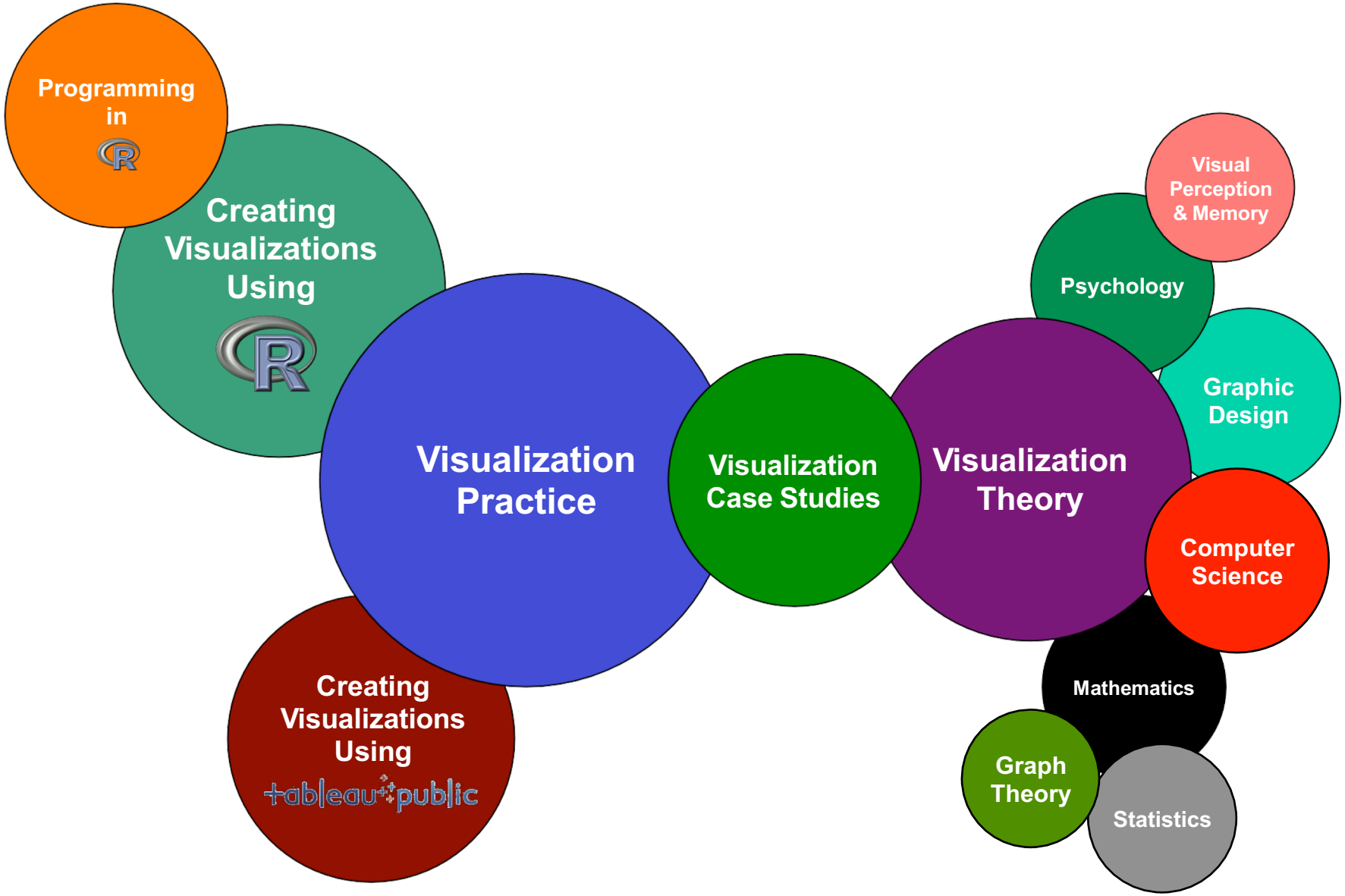
https://www.r-graph-gallery.com/all-graphs

6. Story telling

- Create, deploy and critically analyse visualisations in story telling



Content


Programming in R
Creating Visualizations Using R
Visualization Practice
Creating Visualizations Using tableau public
Visualization Case Studies
Visualization Theory
Visual Perception & Memory
Psychology
Graphic Design
Computer Science
Mathematics
Graph Theory
Statistics

Books

- Designing Data Visualisations – Noah Iliinsky, Julie Steele (2011)
- Visualise this – Nathan Yau (2011)
- Visualising Data – Ben Fry (2008)

Also worth looking up:

- The Visual Display of Quantitative Information – Ed Tufte (1983)
- Semiology of graphics – Jacques Bertin (1967)
- Andy Kirk – Data Visualisation (2016)

Remember!

- Well-designed visual representations can replace cognitive calculations with simple perceptual inferences and improve comprehension, memory, and decision making

- By making data more accessible and appealing, visual representations may also help engage more diverse audiences in exploration and analysis

- The challenge is to create effective and engaging visualizations that are appropriate to the data, the audience and the message

Thanks To

- Cathy Ennis, Marisa Llorens-Salvador, John McAuley, Colman McMahon and Brian Mac Namee for an earlier version of these lecture notes