Communicating Complex Data: Creating Understanding Through Visualization

Marc-Oliver Wright
MT(ASCP), MS, CIC, FAPIC

Clinical Science Liaison, PDI Healthcare
marc-oliver.wright@pdihc.com

Disclosure

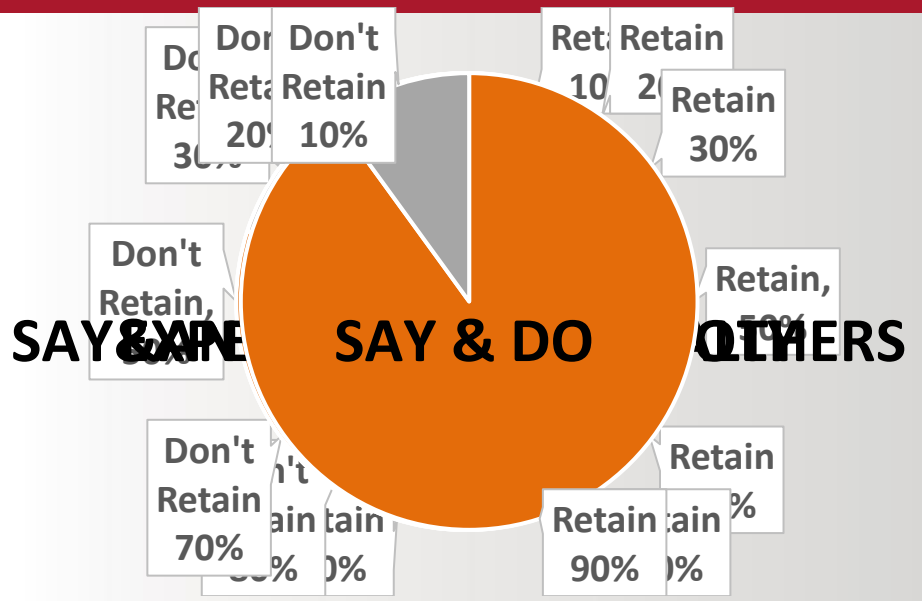
- I am an employee of the clinical team of PDI Healthcare. The content of this presentation is not representative of the views of PDI or its ownership.
- Presentation will incorporate best practices from a variety of information sources that bridge medical disciplines.
- There will be NO discussion of any PDI products and/or solutions in accordance with CE Requirements.
- The content of this session was previously presented at APIC 2022.

Objectives

Upon completion of this program, the participant will be able to:

- Describe the importance of effective communication of data and how design of graphical representations can help or hinder comprehension using examples.
- List the nine principles of effective graphical display of data.
- Deploy skills of graphical integrity, maximizing data-ink and avoiding chart junk with visual aesthetics to convey data to the broadest possible audience.

Acknowledgements



Adults **retain**:
10% of what they read
20% of what they hear
30% of what they see
50% of what they see and hear
70% of what they say/discuss with others
80% of what they experience personally
90% of what they say and do [1]

Retention ≠ Comprehension

1+x

P=0.004

2.02 per 1,000 line days

1,478 catheter days

3 deaths

SIR=1.25

1.17%

T R U T H

2.35 per 100 admissions

1.75 FTEs

4 CLABSIs

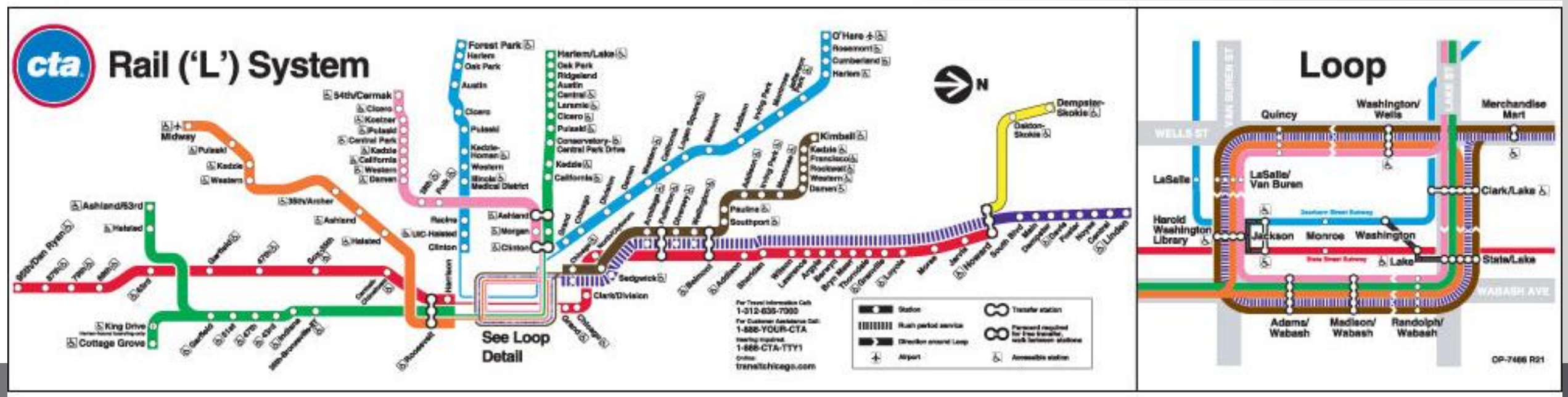
500 RLU

1.10 per 10,000 patient days

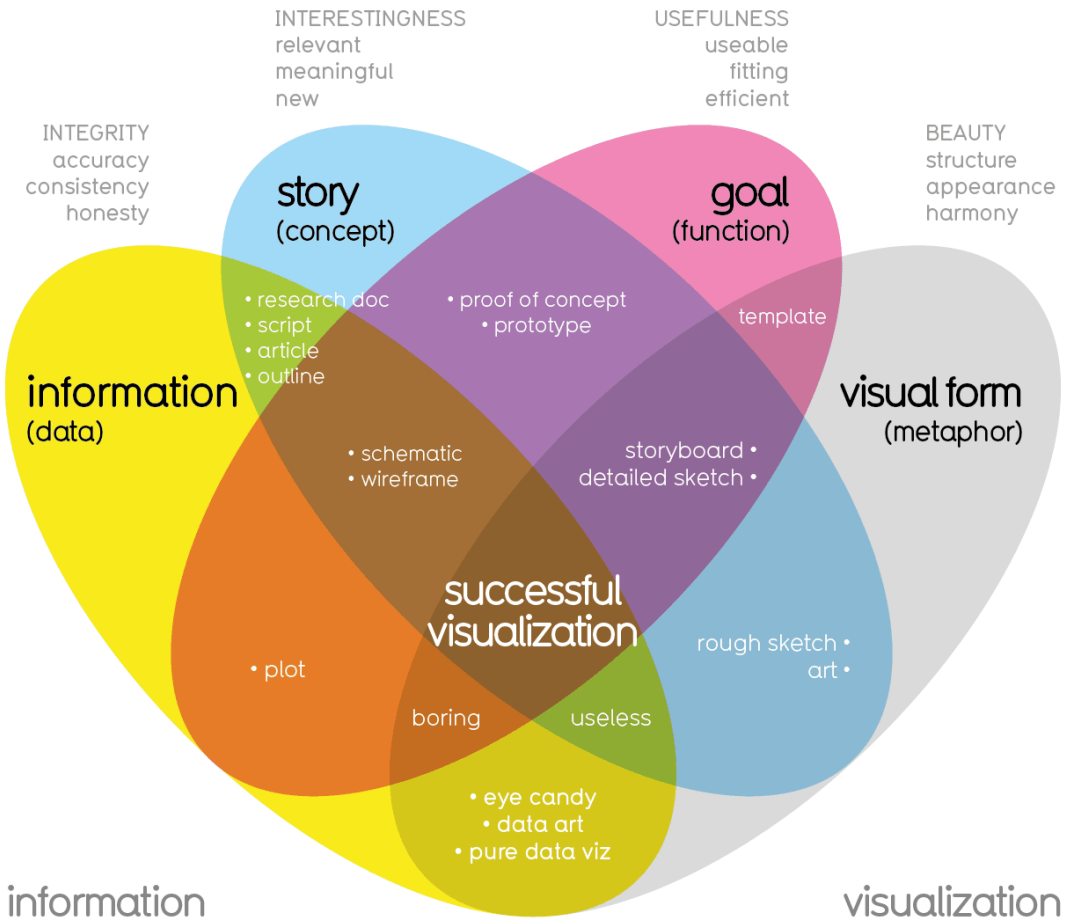
2 employee exposures

Beginning Your Data Journey

- *Know your audience*
- *Develop a summary statement*
- *Do a test run*



What Makes a Good Visualization?



information
the art of journalism
structuring & harmonising information

visualization
the art of design
structuring & harmonising visuals

David McCandless
InformationisBeautiful.net

find out more
bit.ly/KIB_Books

Nine Principles

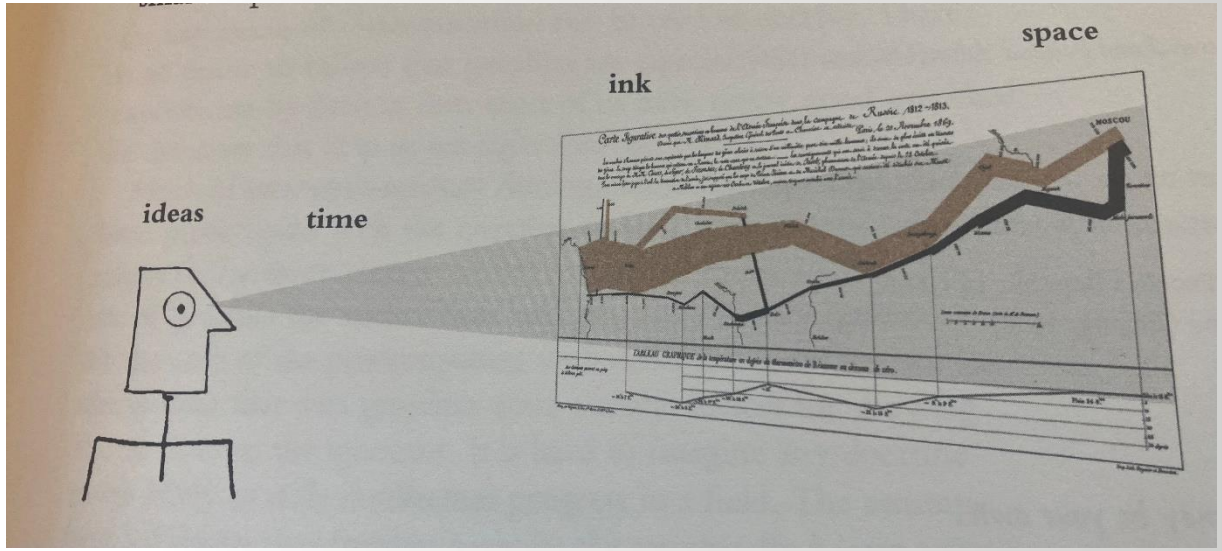
- *These 9 elements are the keys to effective graphical display of data.*

  Balance
  Emphasis
  Movement
  Pattern
  Repetition
  Proportion
  Rhythm
  Variety
  Unity

https://www.getty.edu/education/teachers/building_lessons/principles_design.pdf

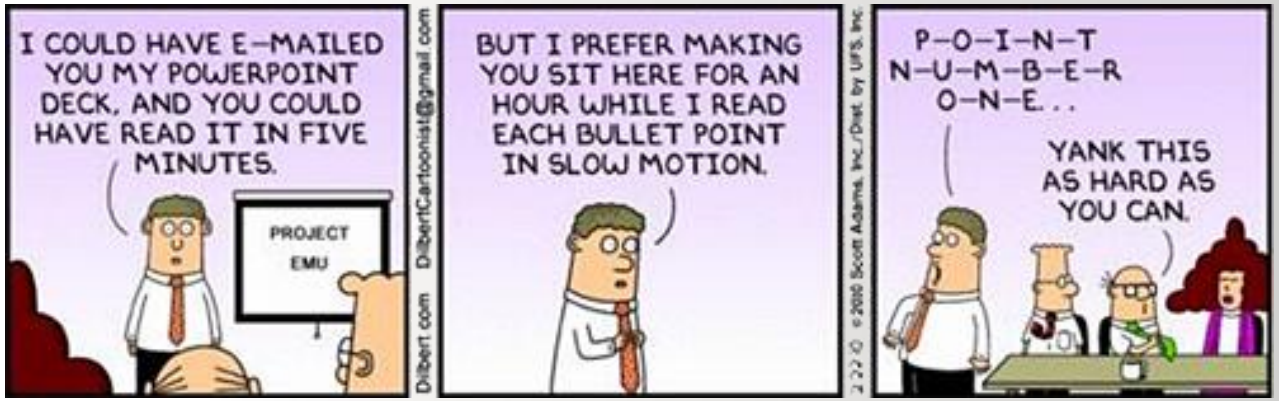
Tufte's Principles of Graphical Excellence

- *Well designed, presentation of interesting data (substance, statistics and design)*
- *Complex ideas shared clearly with precision and efficiency*
- *Gives the viewer*
  - *The most ideas quickly*
  - *With the least ink*
  - *In the smallest space*
- *Multivariate*
- *Tells the truth (about the data)*



E Tufte in The Visual Display of Quantitative Information, 2nd Edition. Graphics Press, Cheshire, Connecticut, 2001 page 51

THIS IS REALLY IMPORTANT



This is less important

You can't read this

You also may not be able to read this.

Excellence in graphical presentation depends on

1. Choosing the best medium for presenting the information
2. Designing the components of the graph in a way that communicates the information as clearly and accurately as possible. [1]

https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/datapresentation/index.html.

Colorful Comments-Emphasis

Color can be a tool to communicate; but we make LOTS of errors

- Good vs. Bad (thematic mistakes-avoid stereotypes)
- Good vs. Bad (daltonism 8%♂/0.5%♀)
- Good vs. Bad (fail to distinguish each other)
- Good vs. Bad (fail to distinguish background)
- Good vs. Better vs. Best (fail to leverage gradient) Good vs. Better vs. **Best**
- Texture can be an option

How We Present Data

Raw Data
- Line lists, when details are needed

Summary Tables
- 2 X 2s
- Characteristics summarized
- Statistics

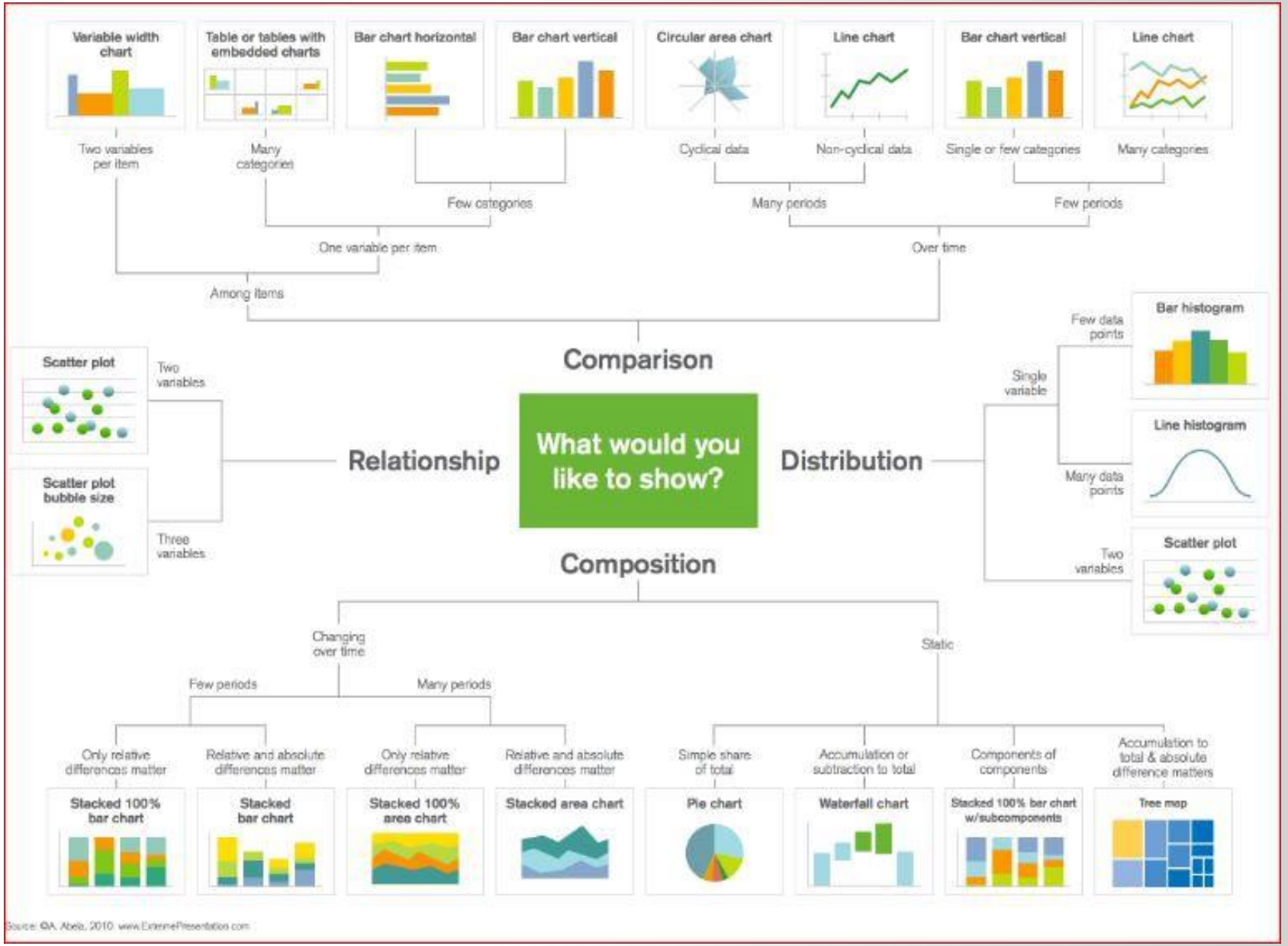
Charts/Graphs
- Pie, bar, SPC, linear, venn diagrams, heat maps

Visual Abstracts & Storytelling

Video
- Sometimes numbers don’t tell the story as well

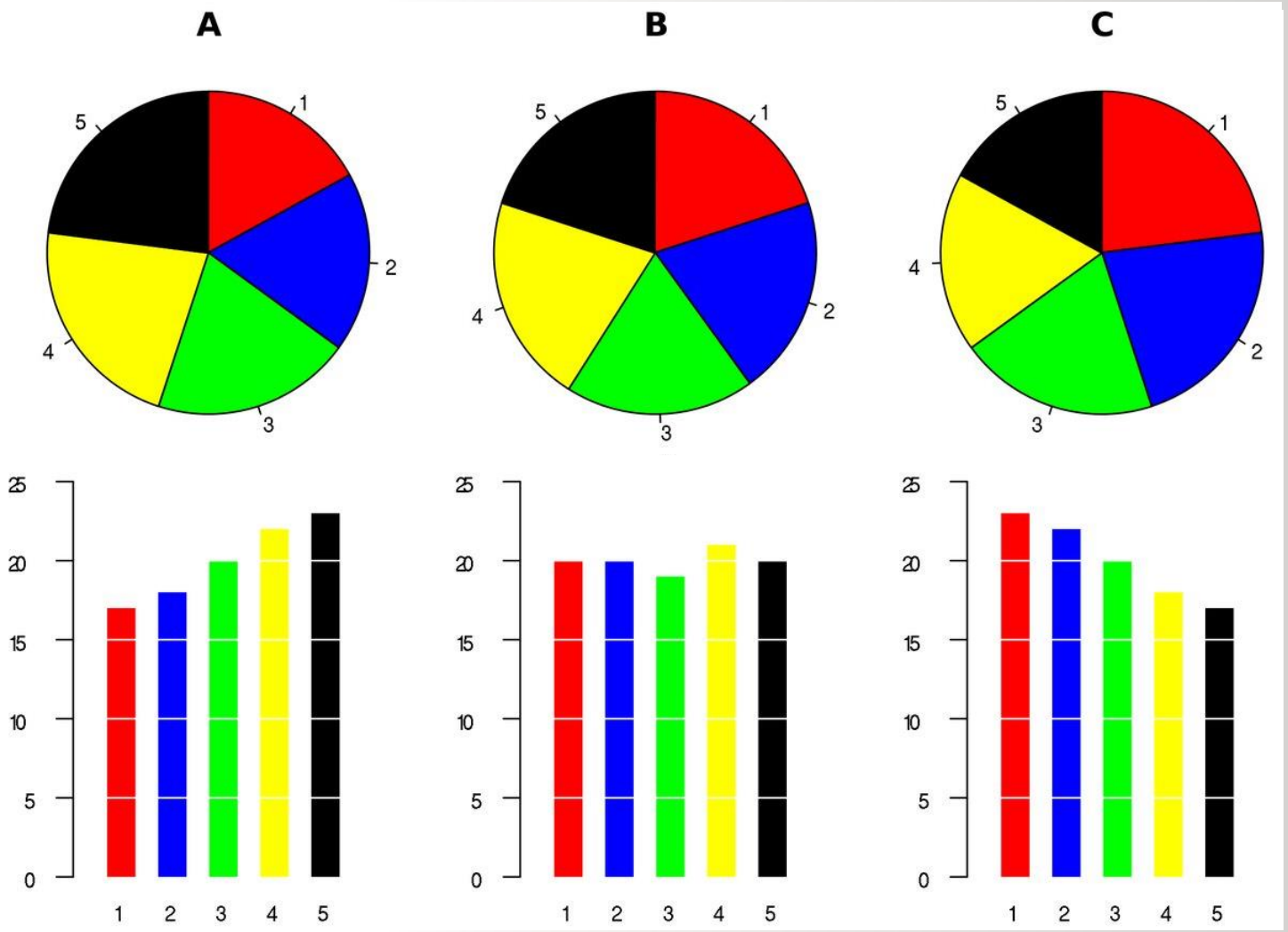
	Tables	Pie charts	Boxplots	Bar charts	Line graphs	Scatterplots
						
Strengths	All data are provided in one place that users can access easily	Useful for depicting a part-to-whole relationship	Help summarise descriptive statistics (median, quartiles, outliers) Useful for depicting the range and distribution of data	Bars can be easily and clearly compared side by side Are versatile and can be used for a number of data types Are widely used and easily understood	Useful for depicting trends and/or predictions over time, where small changes in the slope of the line are easy to see	Useful for depicting relationships between continuous variables Can fit a line to indicate the strength of the relationship Easy to see any unusual observations
Limitations	Can be too detailed and cluttered Main findings in the data are not immediately obvious	Difficult to read when slices are similar Not suitable when there are many slices, categories are non-mutually exclusive, or do not sum together	Do not show relationships between variables	Less useful for depicting the strength of the relationship between variables	Too many lines can be difficult to read Not suitable or potentially misleading when applied to some types of data (e.g., categorical)	Cannot summarise descriptive statistics Not appropriate for some types of data (e.g., categorical)

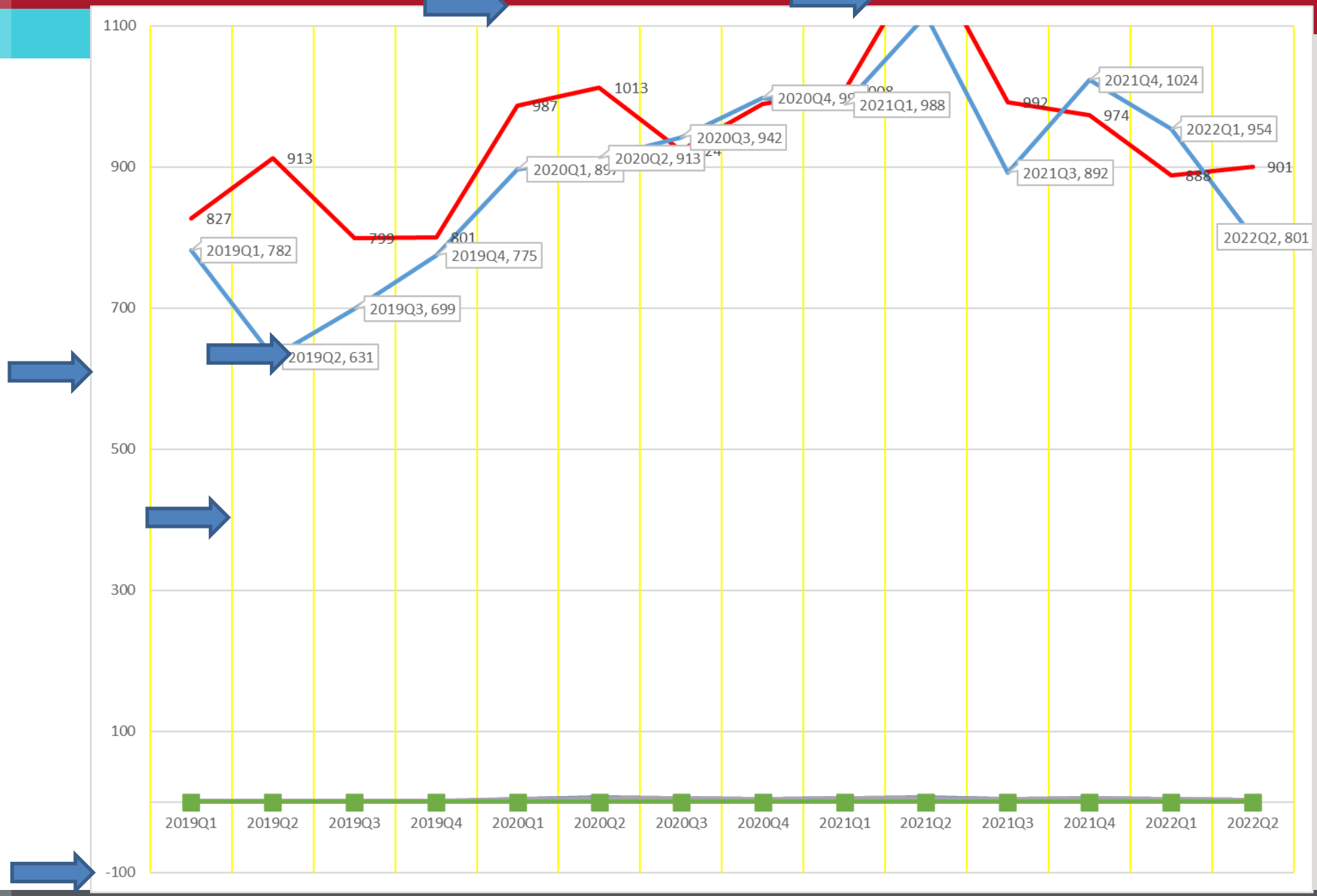
Borg, Zhou and Tull. Exploring the Problem with Data Chapter 3, Fig 4 in The Method DOI: 10.26180/14111366.v1

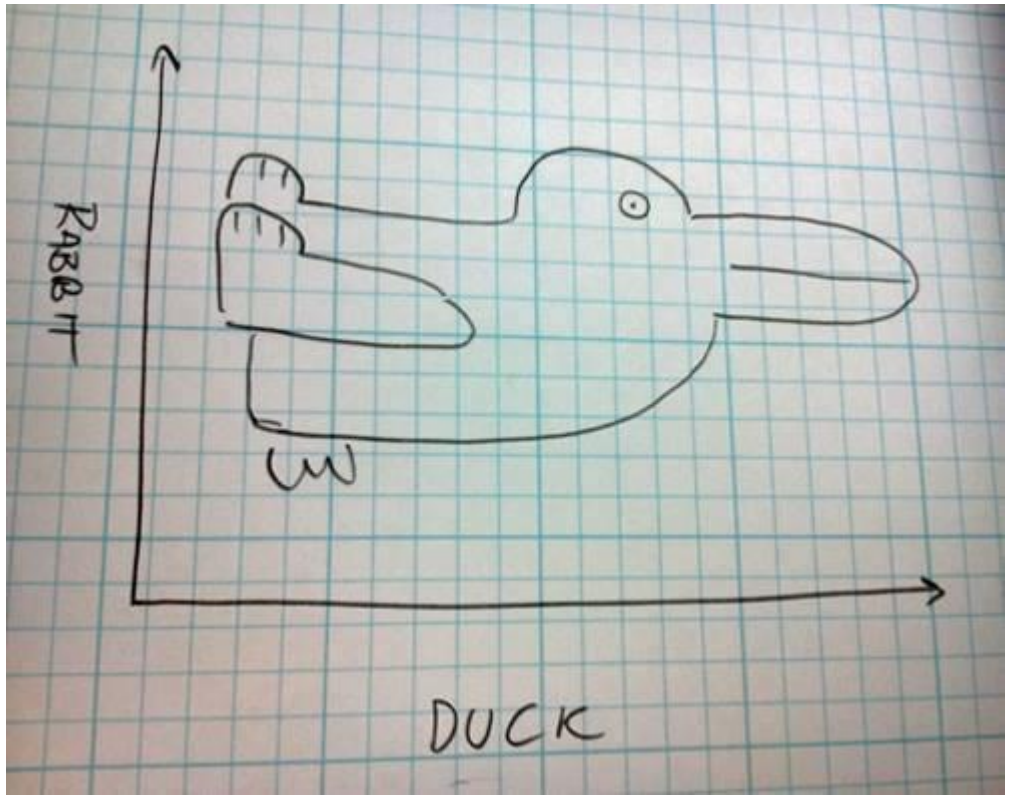


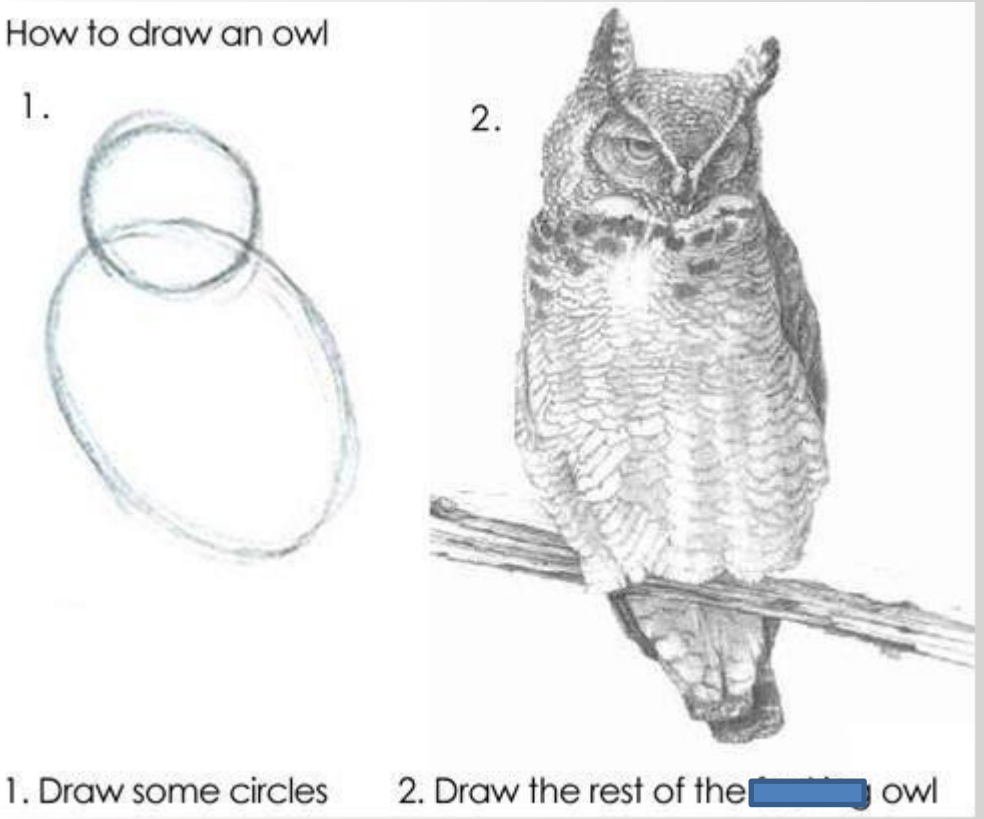
More detailed but similar version available here: https://ft-interactive.github.io/visual-vocabulary/

Which one of these is not like the others?




1100
900
700
500
300
100
-100
2019Q1
2019Q2
2019Q3
2019Q4
2020Q1
2020Q2
2020Q3
2020Q4
2021Q1
2021Q2
2021Q3
2021Q4
2022Q1
2022Q2
827
913
799
801
987
1013
974
992
888
901
2019Q1, 782
2019Q2, 631
2019Q3, 699
2019Q4, 775
2020Q1, 89
2020Q2, 913
2020Q3, 942
2020Q4, 99
2021Q1, 988
2021Q3, 892
2021Q4, 1024
2022Q1, 954
2022Q2, 801


RABBIT
DUCK


How to draw an owl
1.
2.
1. Draw some circles
2. Draw the rest of the owl

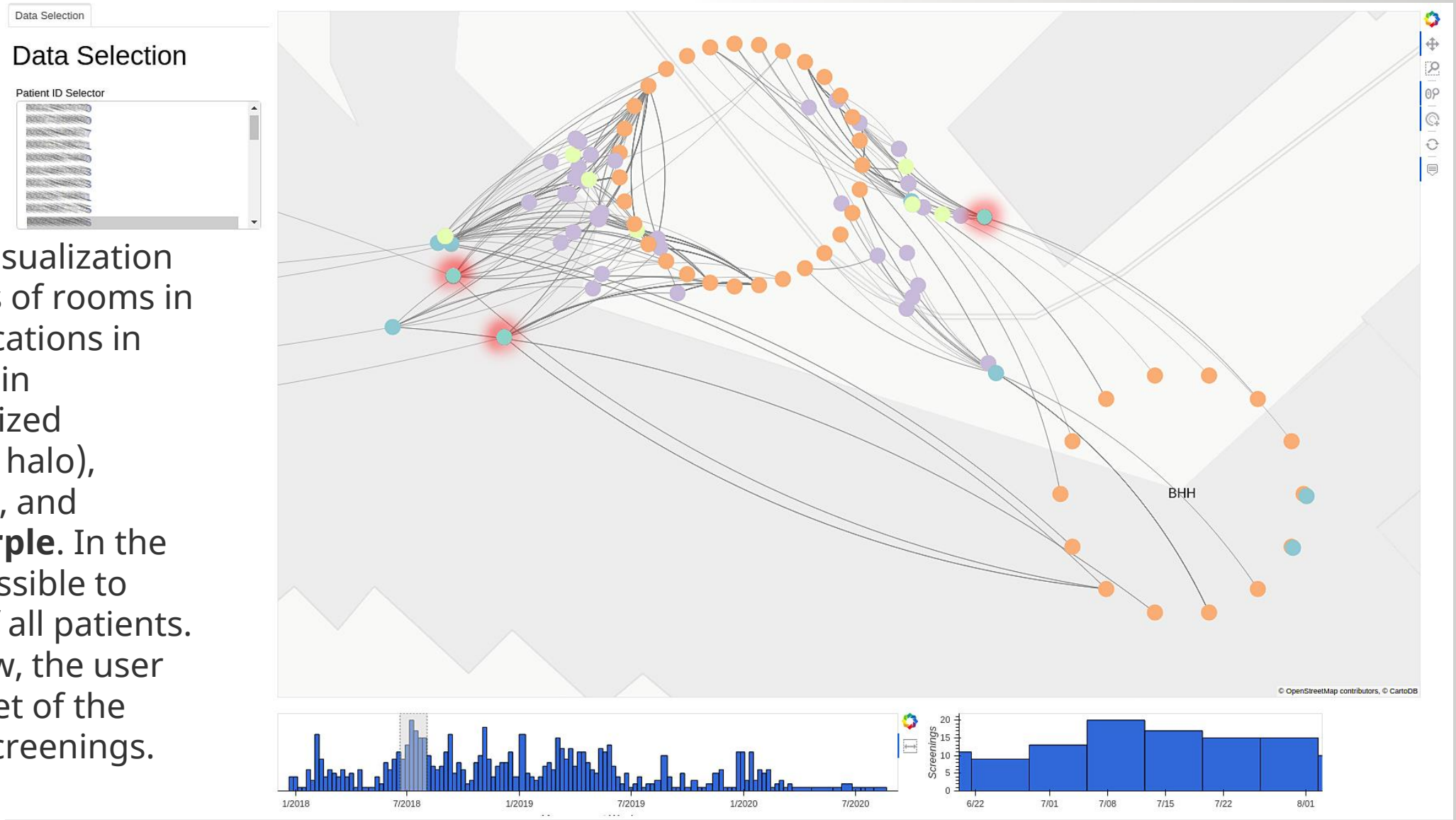
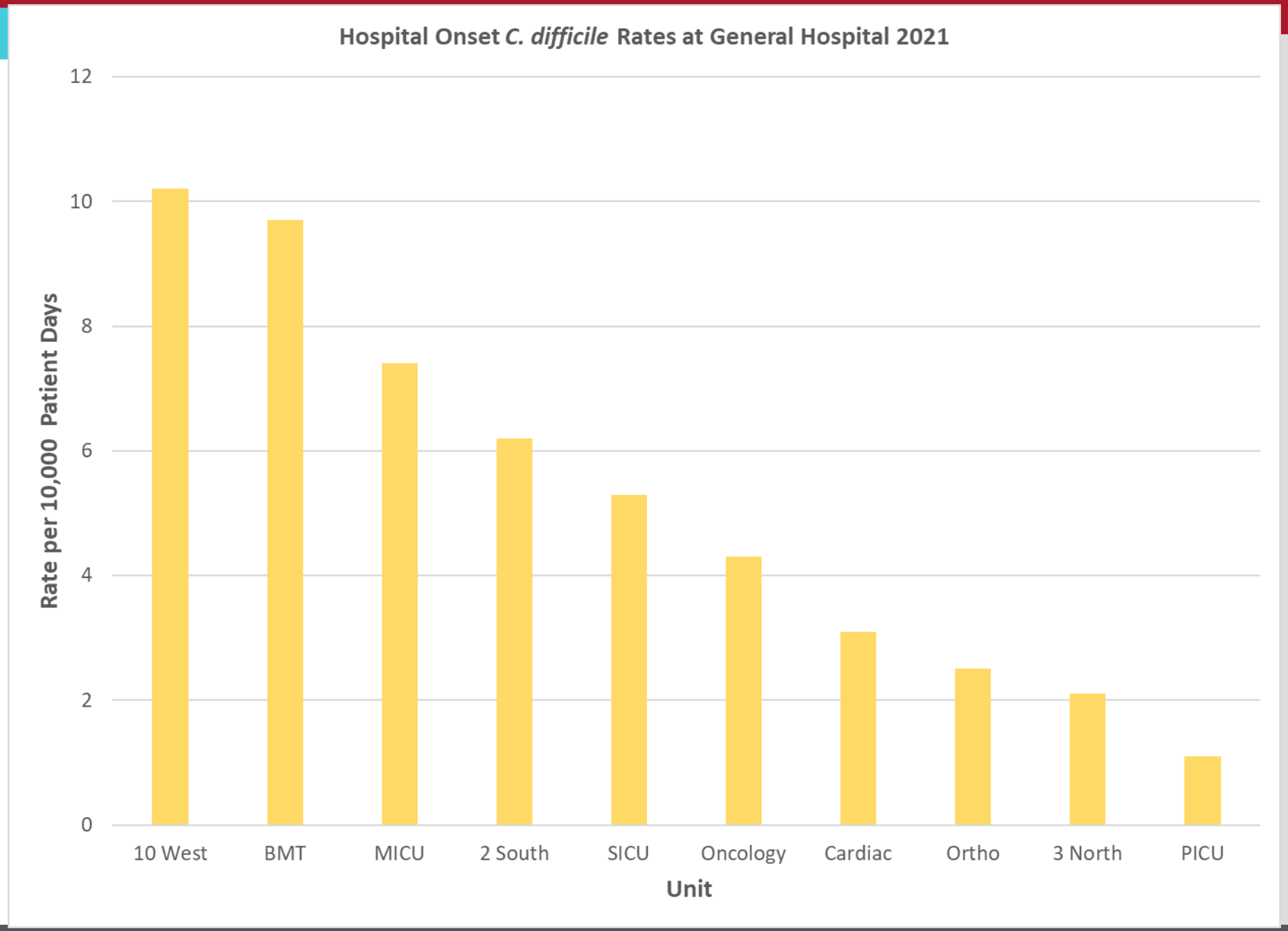


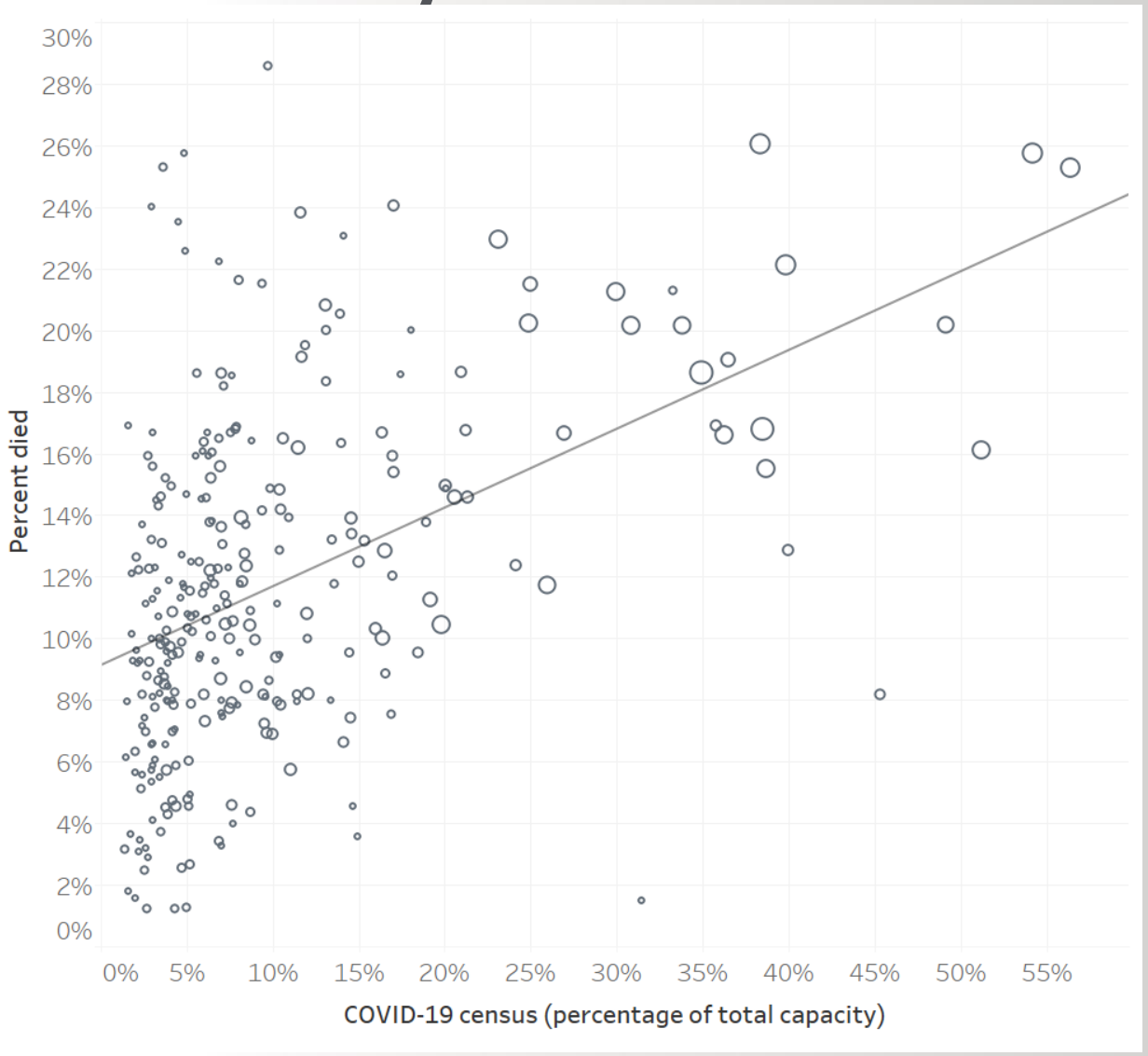
Fig. 1. Example visualization shows collections of rooms in the geospatial locations in **orange**, patients in **turquoise** (colonized patients with **red** halo), devices in **yellow**, and employees in **purple**. In the left panel, it is possible to select a subset of all patients. In the bottom row, the user can select a subset of the timeline of VRE screenings.

Atkinson, A., et al. ICHE 2023 *44*(2), 246-252. doi:10.1017/ice.2022.66


Hospital Onset C. difficile Rates at General Hospital 2021
Rate per 10,000 Patient Days
12
10
8
6
4
2
0
10 West
BMT
MICU
2 South
SICU
Oncology
Cardiac
Ortho
3 North
PICU
Unit

Relationship Between Hospital Mortality & COVID-19 Census

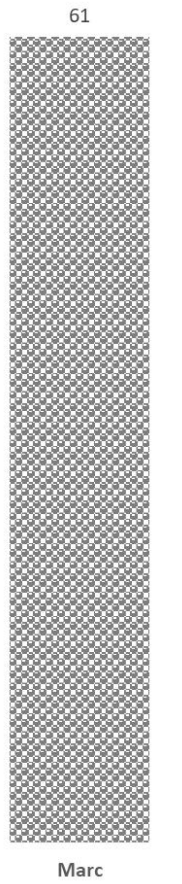
Scatterplot + trendline
Titled, labelled with faint gridlines
Diameter of points
No equation/goodness-of-fit

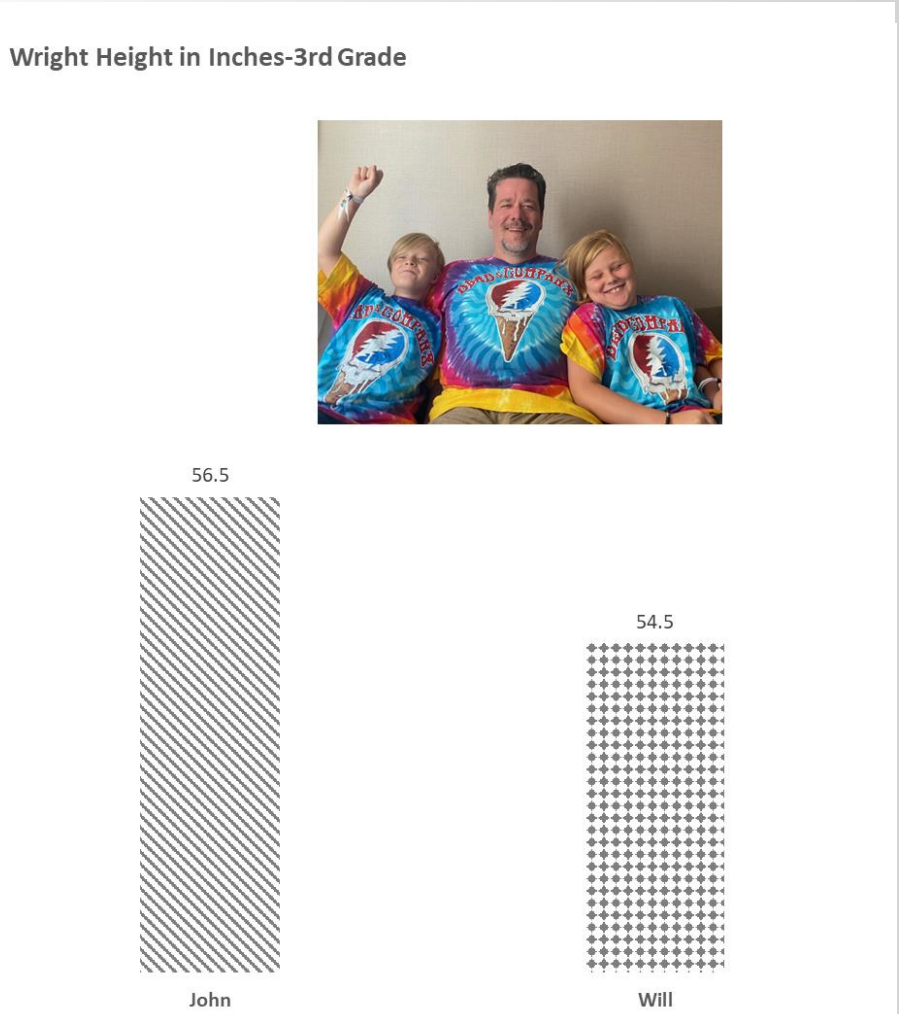


Rules of Data Ink and ChartJunk

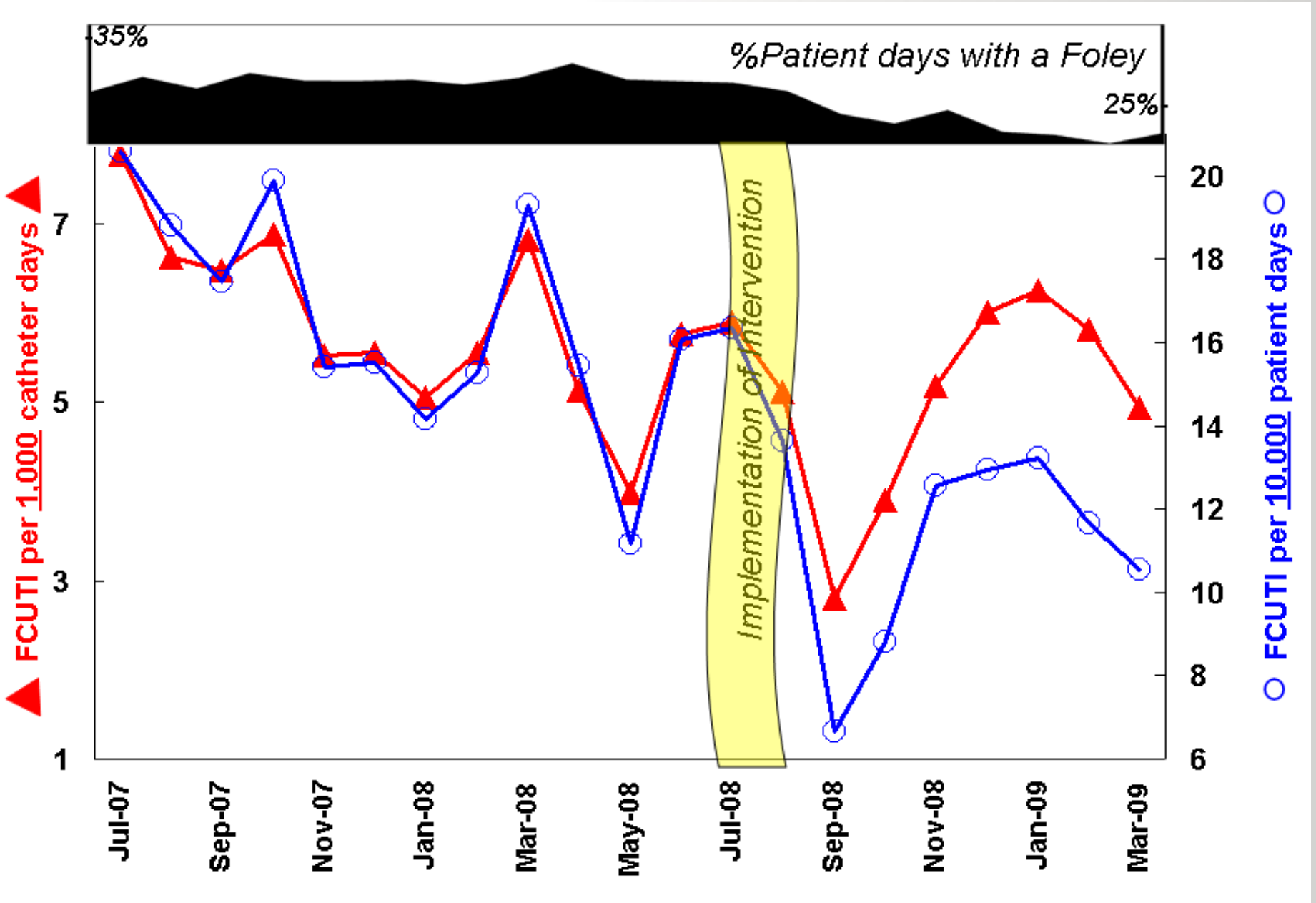
- *Above all else, show data.*
- *Maximize the data-ink ratio.*
- *Erase non-data-ink.*
- *Erase redundant data-ink.*
- *Revise and edit.*

- *Less is actually more*
- *What purpose does this serve?*
- *What will be lost if I remove it?*



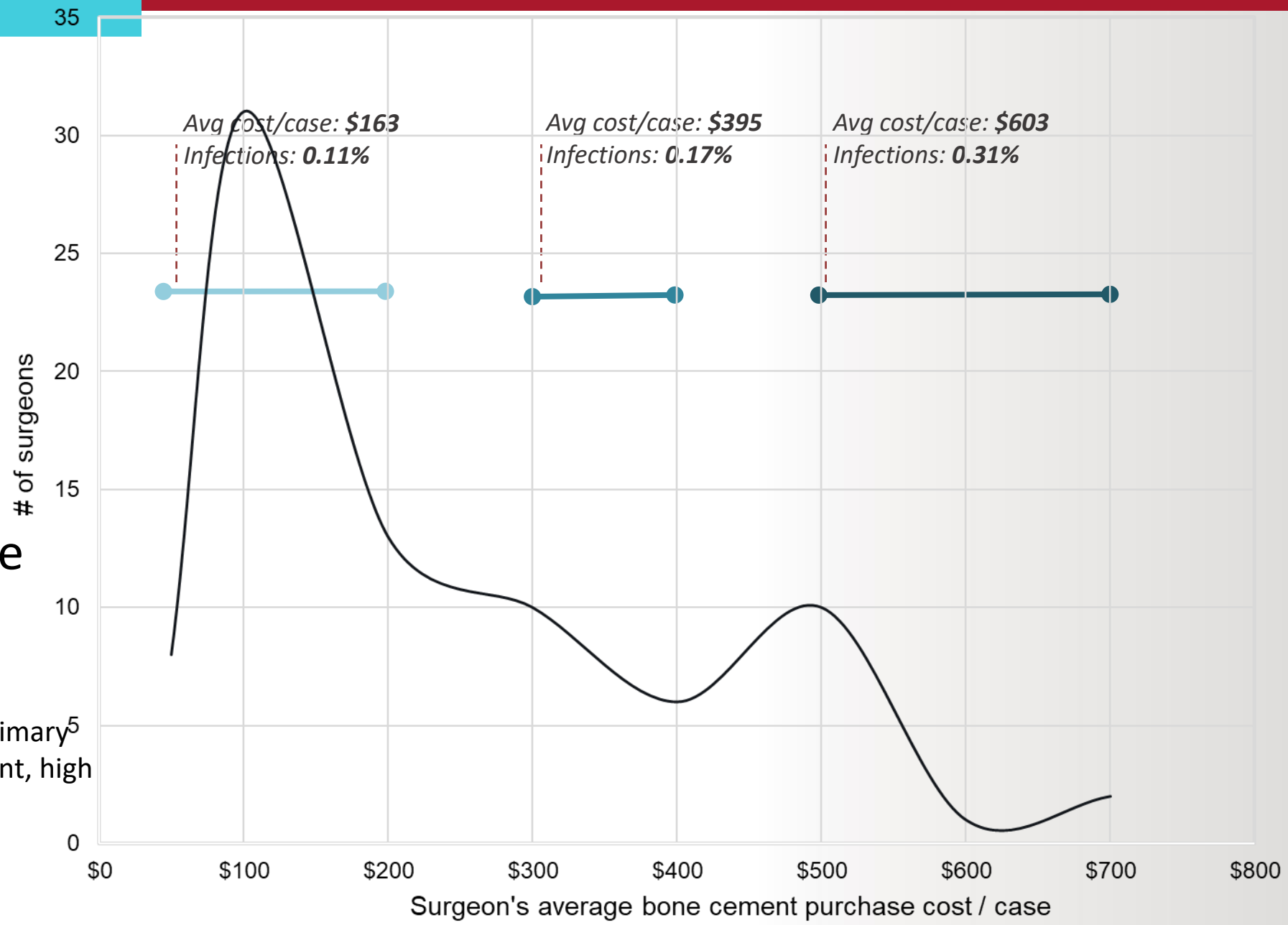


E Tufte. The Visual Display of Quantitative Data Chapter 4-5

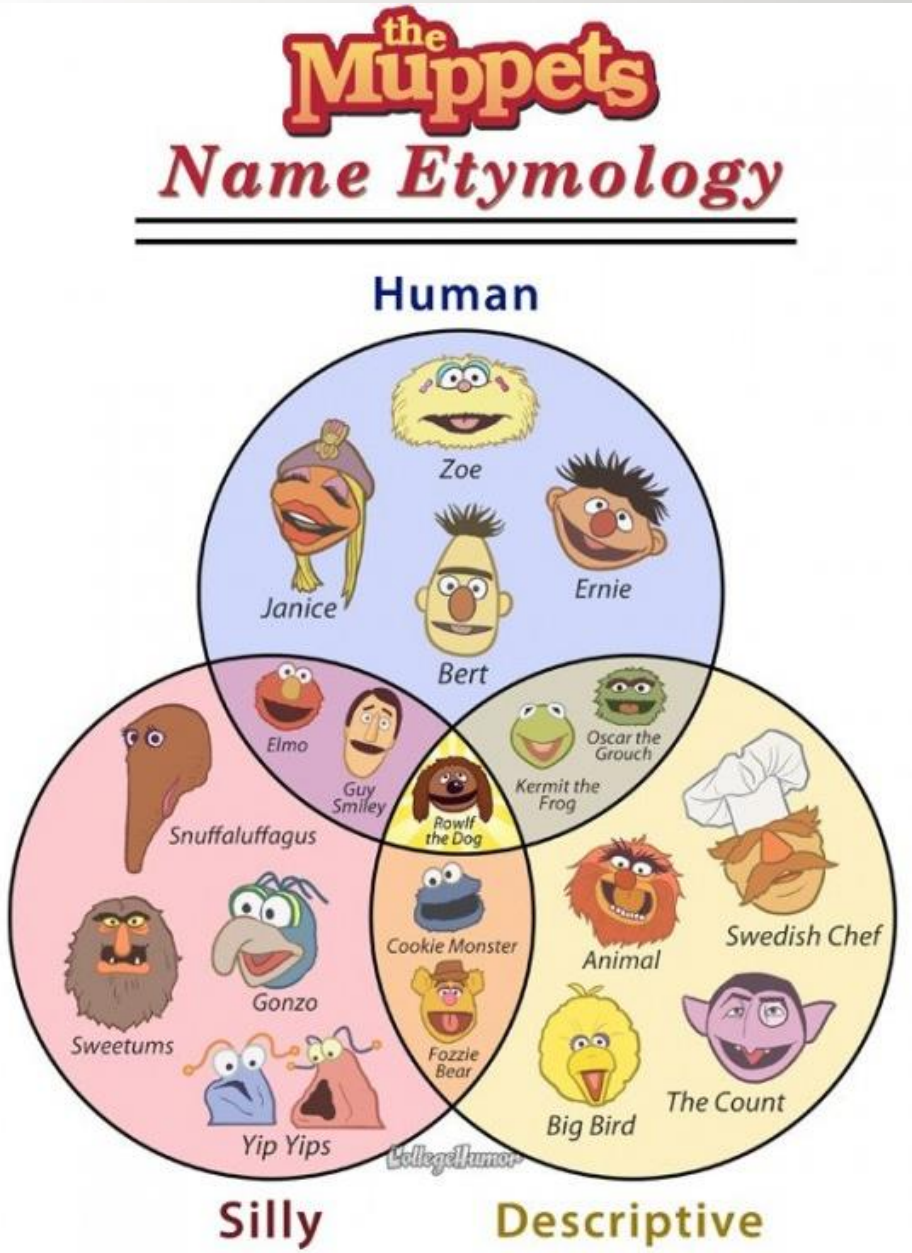

35%
%Patient days with a Foley
25%
Implementation of intervention
FCUTI per 1,000 catheter days
FCUTI per 10,000 patient days
1
3
5
7
6
8
10
12
14
16
18
20
Jul-07
Sep-07
Nov-07
Jan-08
Mar-08
May-08
Jul-08
Sep-08
Nov-08
Jan-09
Mar-09

Average bone cement cost/case

Elective, Unilateral, Primary Total Knee Replacement, high volume surgeons
n = 20,000 cases



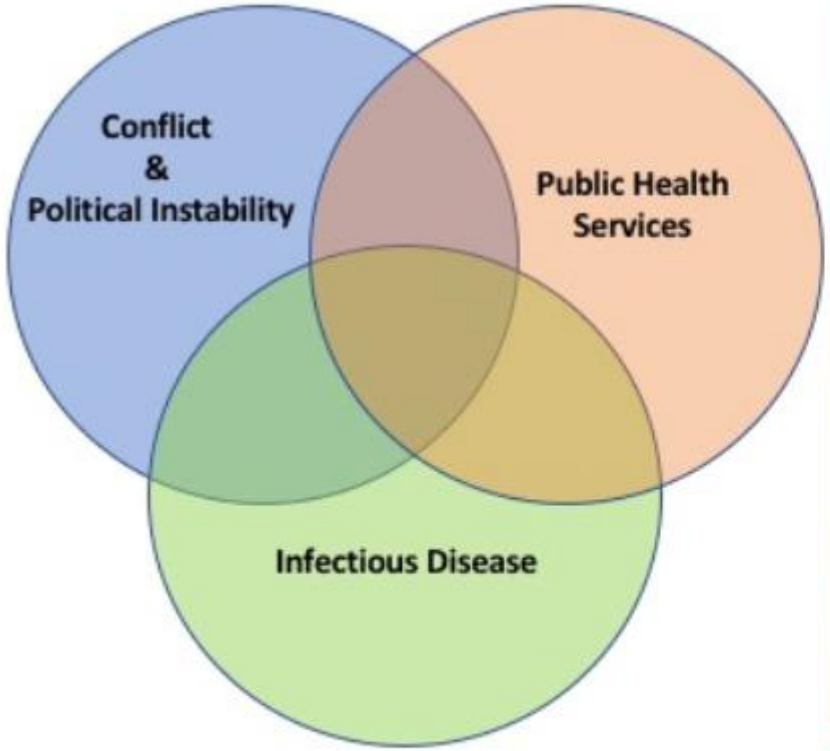
- *Not all are Muppets-some are exclusively Sesame Street*
- *Not properly, spatially placed-*
- *I would argue Ernie is closer to silly than Bert for example-ditto Swedish Chef vs Big Bird*
- *Isn't Guy Smiley also descriptive?*
- *Isn't Swedish Chef the most human?*
- *Who knows a human named Elmo?*
- *We don't call him "The Count" he's just Count*
- *They're Martians, not Yip Yips*
  Rowlff the Dog, is by definition, not human



How not to use a Venn diagram

No title

Which overlap and where

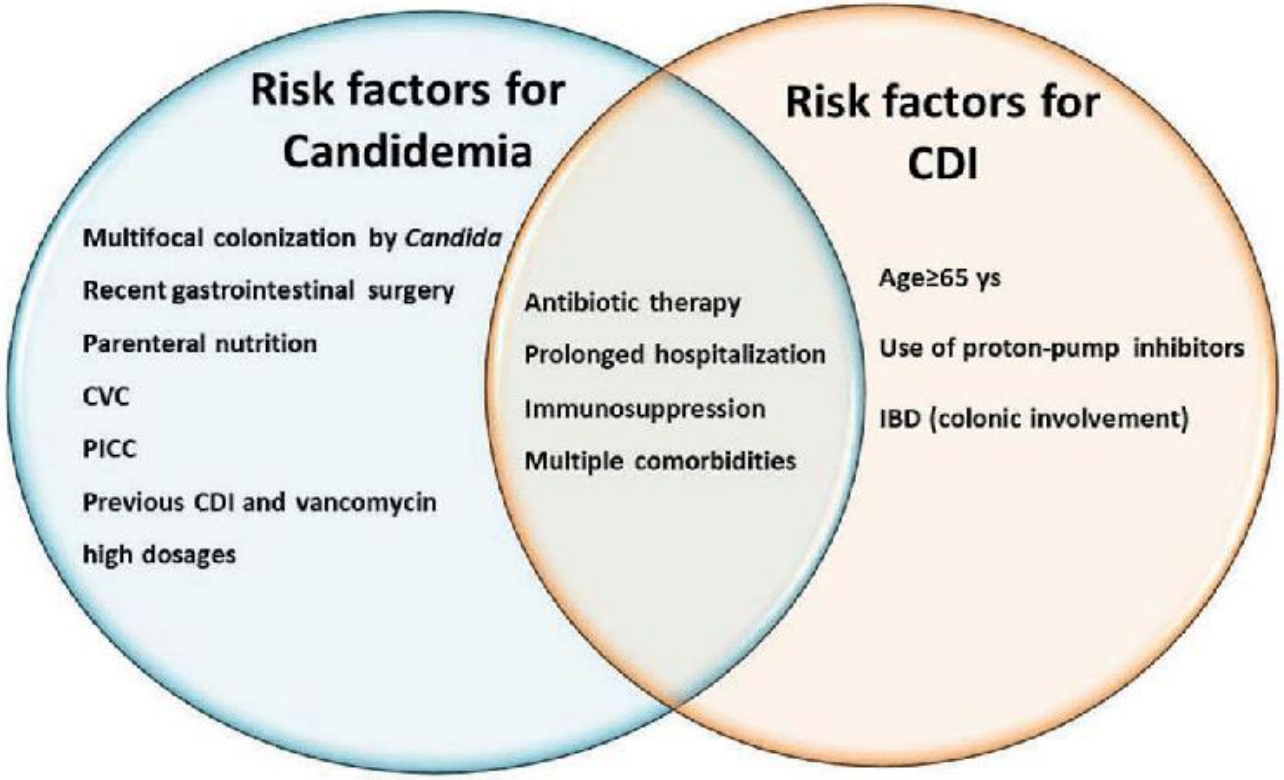


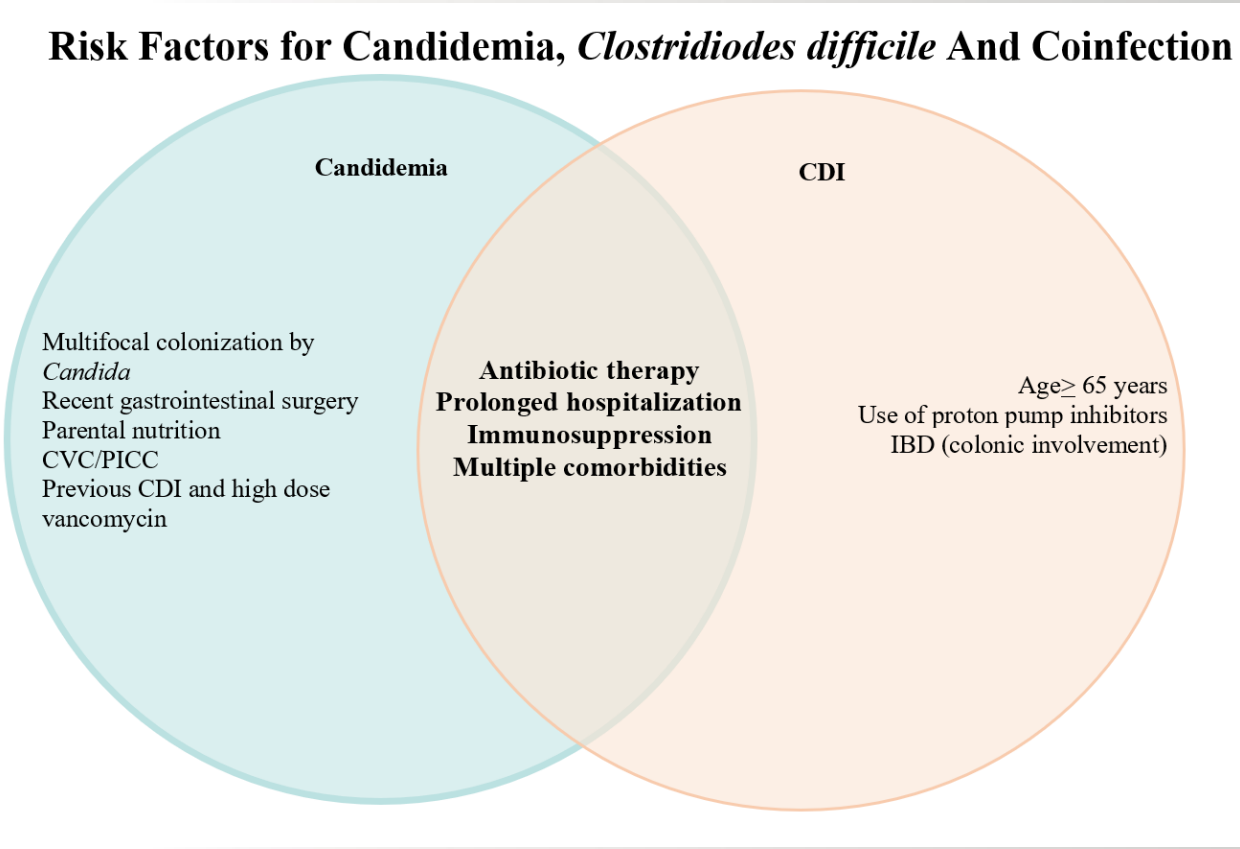
- Insecurity and inaccessible populations
- Supply chain disruption
- Inflation
- Lack of governance, civil servants go without pay
- Internally displaced communities
- Insecure borders

- Breakdown of surveillance systems
- Closed health facilities (due to violence or lack of governance/funding)
- Lack of infection prevention and control measures
- Reduced/limited workforce for health centers
- Poor sanitation, access to clean water
- Food shortages
- Malnutrition
- Stalled vaccine campaigns
- Lack of medication

- Vaccine preventable diseases arise
- Food/water-borne diseases emerge
- Increases in vector-borne diseases
- Re-emerging infectious diseases (cholera, diphtheria, polio, measles)

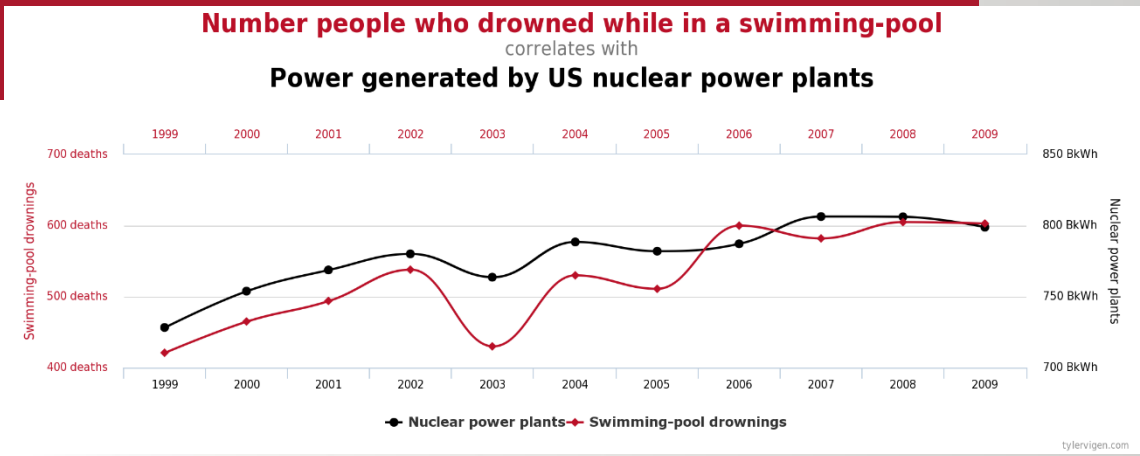
https://ghss.georgetown.edu/idconflict/





•M Falcone Expert Review of Anti-infective Therapy 17(62)
DOI:10.1080/14787210.2019.1608183

Statistics



- *Know your audience*
- *Make sure you are using the right tool*
- *Find a friend*
- *Explain what you are using, why and what it means*
- *Odds ratio: measures the strength of an association between an exposure and an outcome between two groups when exposures, cases and controls are well defined.*

	Developed CDI	Did Not Develop CDI
Admitted to a room of a recently discharged CDI+ patient	22	11
Admitted to a room of a recently discharged CDI- patient	9	28

- *OR = (22X28)/(11X9) = 6.22 – patients admitted to rooms of recently discharged CDI patients are 6.22 times more likely to develop CDI*
- *Beware Correlation <> Causation*

The problem with p-(values)

Some points from the ASA

P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.

Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.

Proper inference requires full reporting and transparency. P-values and related analyses should not be reported selectively

A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.

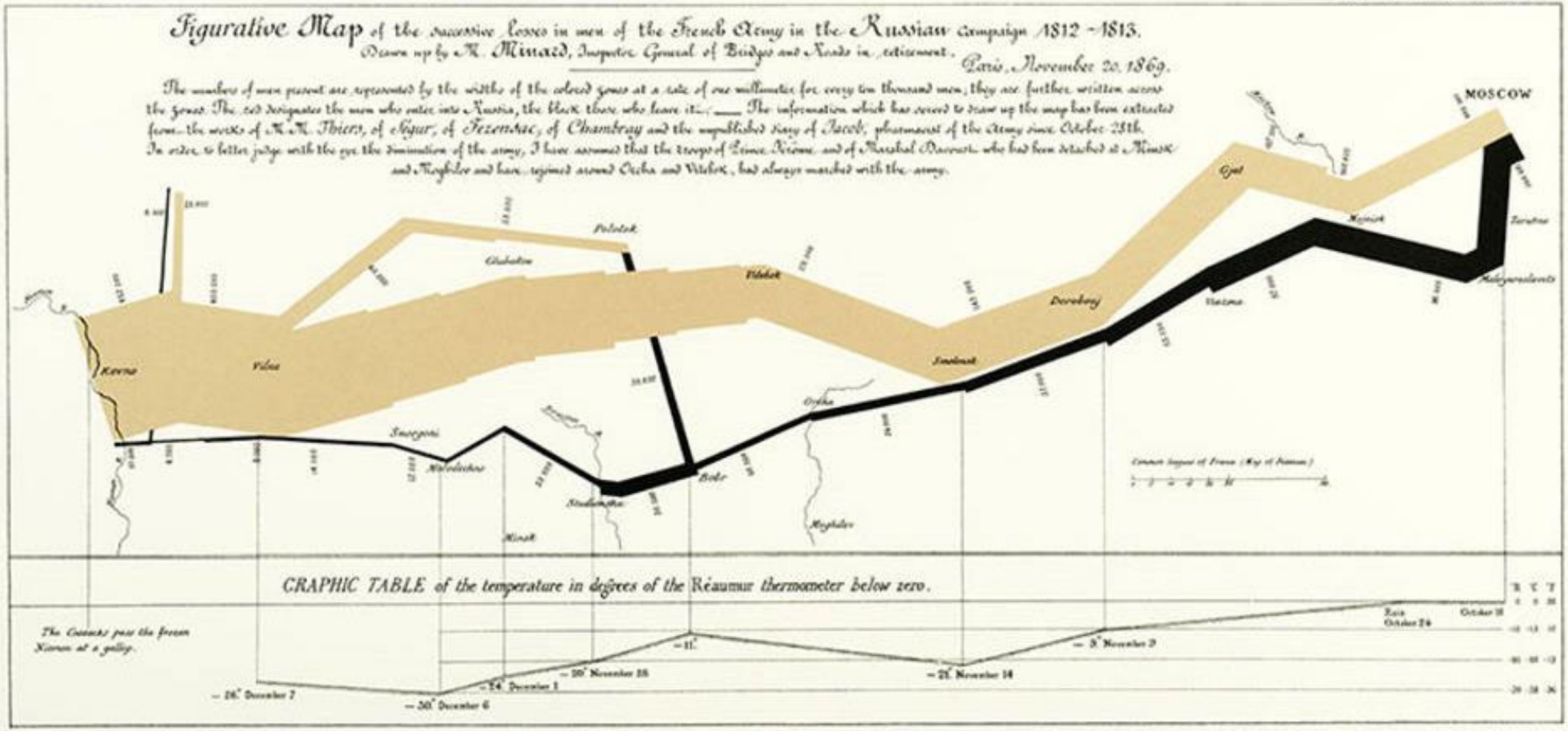
By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis. (Wasserstein and Lazar https://doi.org/10.1080/00031305.2016.1154108)

" If your statistics yield a *P* value of .06, then it's a negative study. If it's .04, then it's much more likely to be considered important... In reality, there's little difference between those findings." (Aaron Carrol https://doi:10.1001/jamahealthforum.2016.0026)

A Most Famous Map

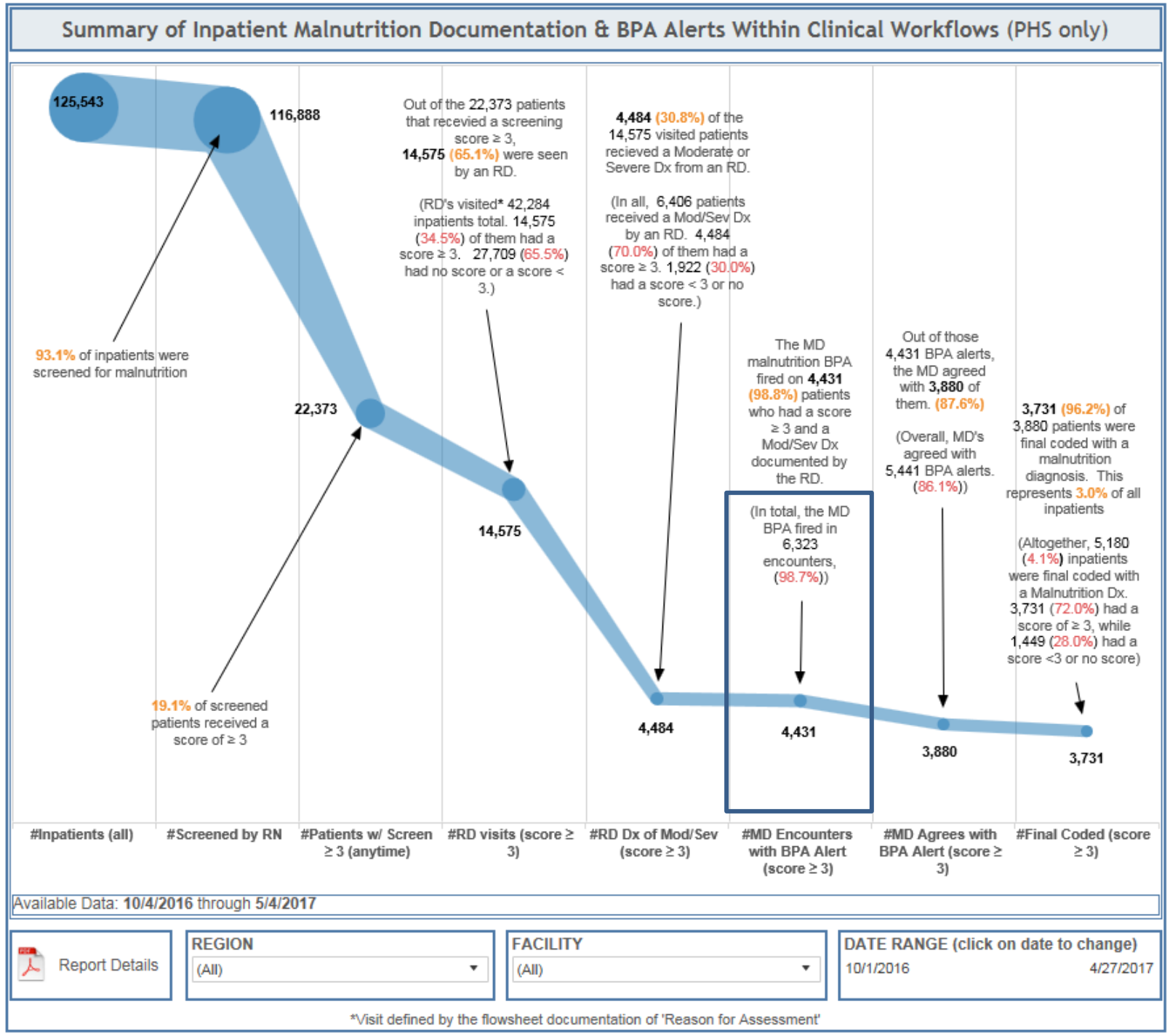
Detection
Assembly
Estimation or Scale

ENGLISH TRANSLATION, 2003



From Edward R. Tufte, *The Visual Display of Quantitative Information* (second edition, 2001) Graphics Press LLC Box 430 Cheshire, Connecticut 06410 www.edwardtufte.com

English translation by Dawn Finley; produced by Elaine Morse. Copyright © 2003 by Graphics Press LLC
For Minard's data sources and a biography of Minard, see www.edwardtufte.com



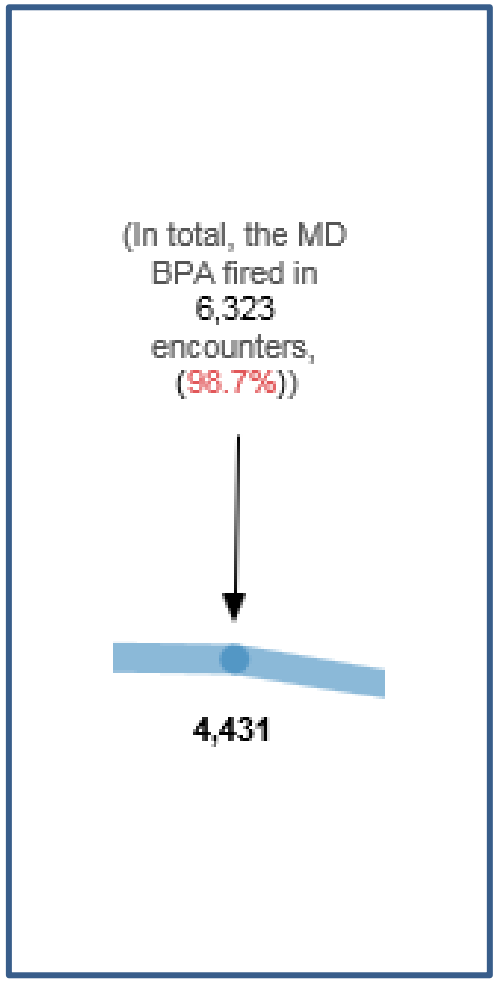
All Regions

RNs doing a great job of screening patients

1/3 of patients who screened positive were not seen by an RD

2/3 of the patients seen by RDs had not screened positive

70% of patients who screened positive didn't have malnutrition



Fired 36,535 times (average 5.8x per encounter)

YouTube Videos as Educational Tools to Promote Hand Hygiene: A Content Analysis

STUDY DATA

400 YouTube videos
"hand hygiene"
"hand hygiene education"

70 analyzed using structured tool

RESULTS

% of videos scored as educationally useful:

55.7%
N = 39

44.3%
N = 31

CONCLUSIONS

Highest Scoring Videos

Attractive
Comprehensive
Persuasive

Recommendation: Guidelines should be used during video development

Lim, K., Kilpatrick C., Storr J., Seale H. *AJIC*
@hollyseale@julesstorr@claireekt
DOI: 10.1016/j.ajic.2018.05.002

APIC Spreading knowledge. Preventing infection.
Association for Professionals in Infection Control and Epidemiology

AJIC American Journal of Infection Control

Are Infection Preventionists and Nurses Engaged in Antibiotic Stewardship?

2017 | Survey of the Corporate Infection Prevention and Control Network | 49% response rate (N = 35)

Role in antibiotic stewardship not well-defined

Current engagement described as "minimal" and "supportive role"

IP Time spent on antibiotic stewardship < 5 hours per month

Gaps in antibiotic stewardship knowledge

1/3 receive no specific training in antibiotic stewardship

Sources of training:

- APIC conference sessions (52%)
- Training given by facility (41%)
- State or local training (31%)
- SHEA training course (7%)

Barriers to Engagement

- Time
- Resources
- Competing priorities
- Lack of role clarity

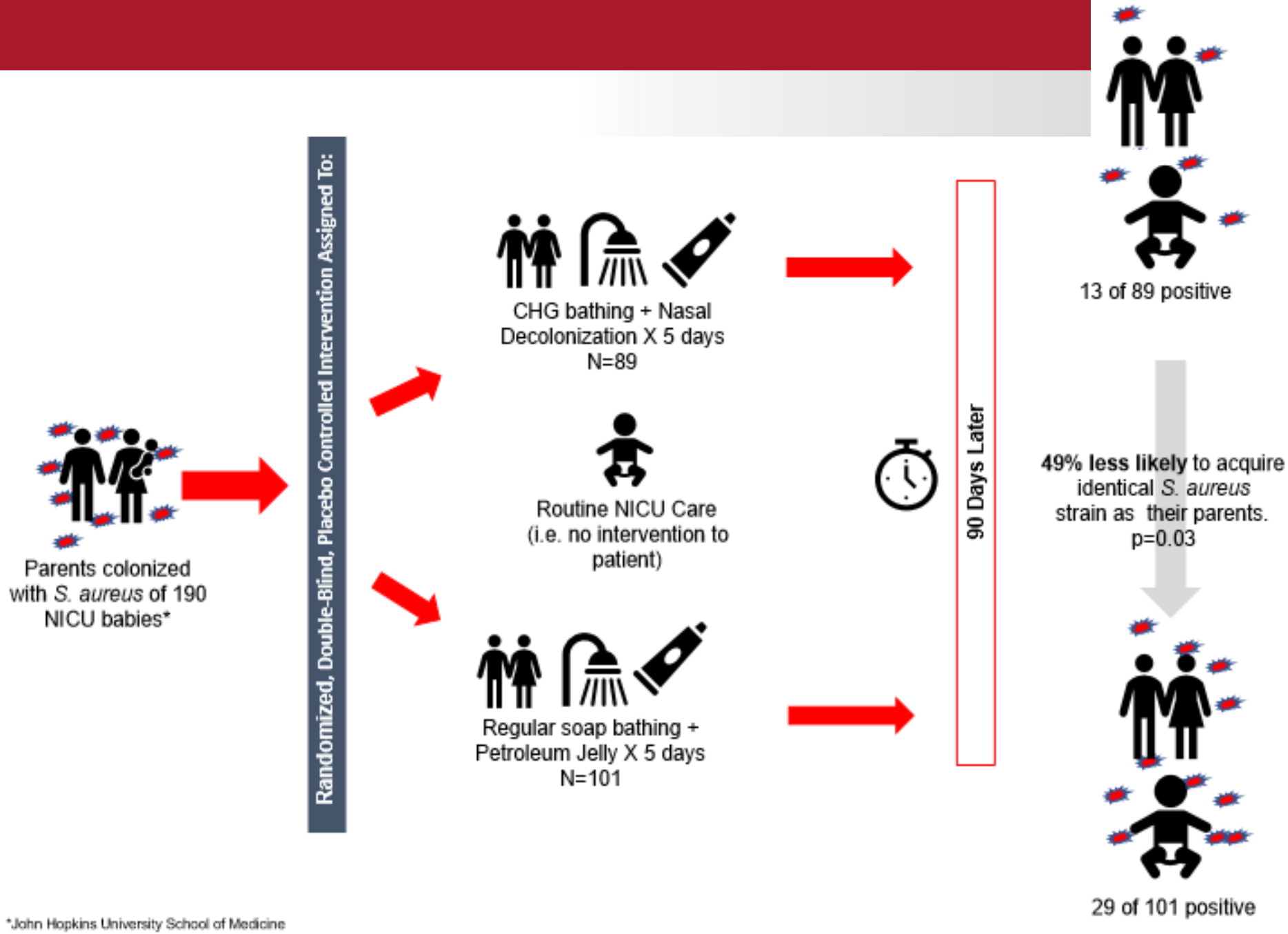
Patient education as a future role in AS

Manning ML & Pogorzelska-Maziarz M. AJIC, In Press.

APIC Spreading knowledge. Preventing infection.
Association for Professionals in Infection Control and Epidemiology

AJIC American Journal of Infection Control

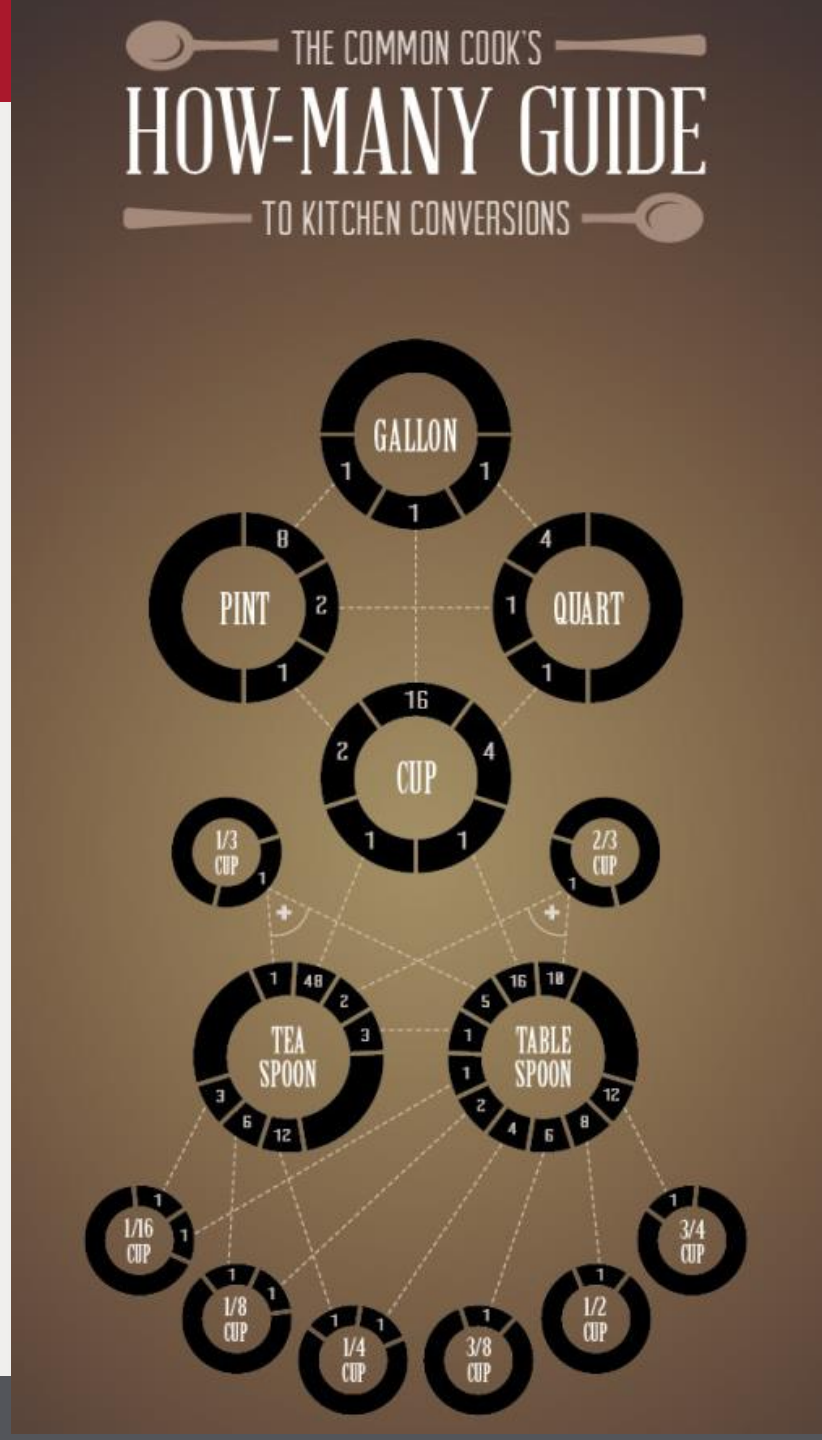
Effect of Treating Parents Colonized With Staphylococcus aureus on Transmission to Neonates in the Intensive Care Unit: A Randomized Clinical Trial



*John Hopkins University School of Medicine

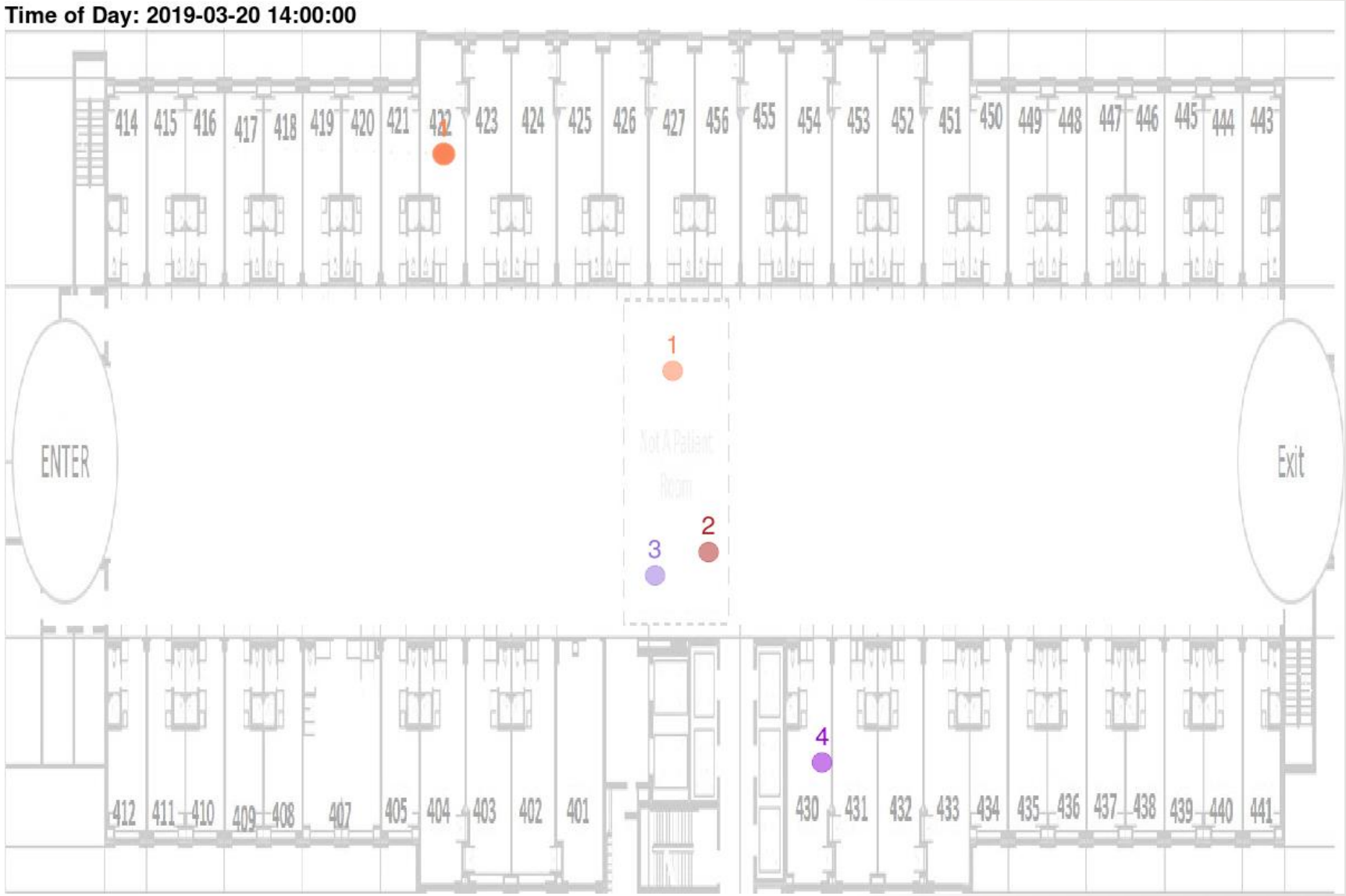
Milstone et al in JAMA 2019 Dec 30;323(4):319-328. doi: 10.1001 avaialble at: https://pubmed.ncbi.nlm.nih.gov/31886828/

Infographics Used in Non-Traditional Ways



http://sblattindesign.files.wordpress.com/2012/09/how-many_guide_infographic1.jpg

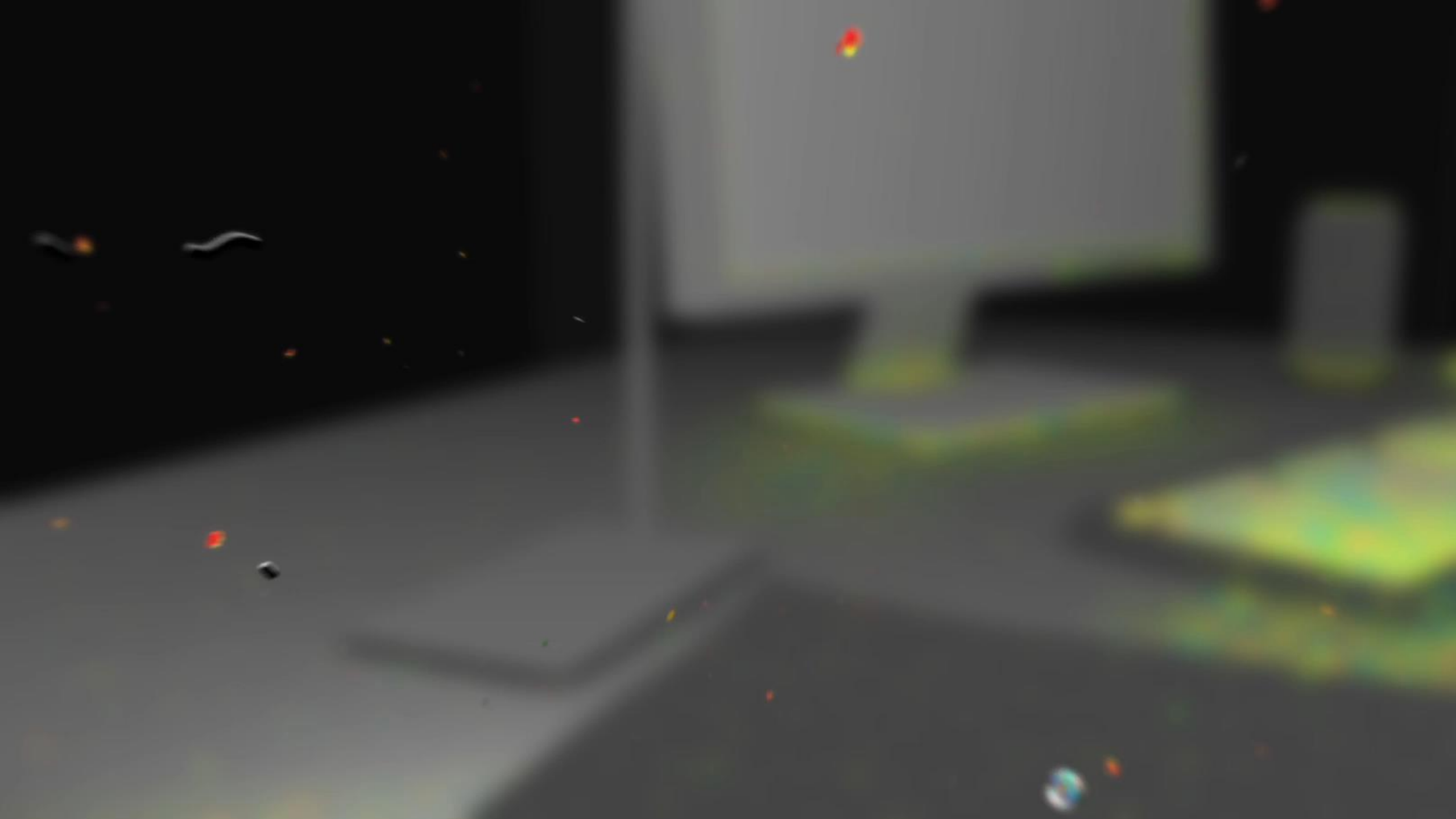
Time of Day: 2019-03-20 14:00:00



1, 2 = NAC
3, 4 = RN

The Value of Video

Video courtesy of Jack Gilbert, PhD Professor in Pediatrics and the Scripps Institution of Oceanography and the University of California, San Diego. Originally shown at ICEID 2018



Summary

- The participant shall be able to list the nine principles of effective graphical display of data.
  - Balance, Emphasis, Movement, Pattern, Repetition, Proportion, Rhythm, Variety, Unity

- The participant shall be able to deploy skills of graphical integrity, maximizing data-ink and avoiding chart junk with visual aesthetics to convey data to the broadest possible audience
  - Avoid textures, redundant information, unnecessary background images, unnecessary backgrounds, gridlines. For every element included in the design ask what purpose does it serve and what would be lost if removed.

Summary

- The participant shall be able to describe the importance of effective communication of data and how design of graphical representations can help or hinder comprehension using examples.
    - Understand your data
    - Write your story first
    - Know your audience
    - Choose the most effective design
    - Be creative
    - Avoid chartjunk, maximize data ink and white space: Simpler is better
    - Do a test run
    - Have fun!

Thank You!