Introduction to information visualizations

CS698Y – Human AI Interaction

Logistics

- Friday (tomorrow) → Attendance is compulsory
- Next assignment coming soon:
  - Conduct a user evaluation of your interface, based on assignment-6
  - Pick a method that suits you → user study (N=3 participants), heuristics evaluation (on your own, for all heuristics, systematically)
  - Have results done, for the final presentations.
  - Final presentations on: 22/23Nov -- report due by 21st.
- Quiz→ 8th November, Saturday.

So far..

- User aspects:
  - Principles
  - Design Processes
  - Explainability
  - Evaluation Techniques
  - Ethics in human studies
- Today & tomorrow → information visualizations

What is visualization?

- Representing data graphically using charts, graphs, and maps to find patterns, trends, and insights that are hard to see in raw data
- Condenses a lot of individual data points, so dense.
- Very common when building ML systems.
- Useful for:
  - Deriving insights
  - Improve communication
  - Helps make decisions

What all do we visualize?

- Data quality
- Model performance
- Explanations

What makes good visualizations

- Accurate
- Clear
- Aesthetically pleasing
- Avoids misleading / miscommunication
- Not too hard to understand
- Hard to do right, but easy to do good enough!

How do you go about visualizing stuff?

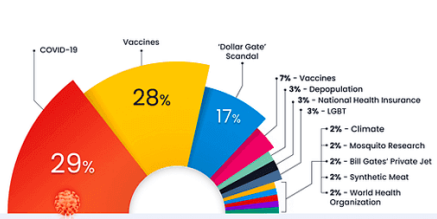
- Here, stuff = numerical/text data, useful for ML
- Start with purpose
- Know your user
- Keep it simple / avoid clutter / too much info. in one visual
- Choice of visualization
- Provide context – in terms of source of data, model parameters, etc. not just labels and titles.
- Use design elements (colors, labels, icons) carefully.

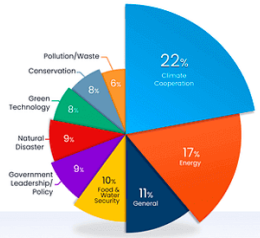
What kinds of visualizations?

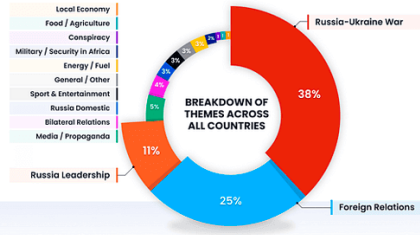
- Plots
  - Scatters, lines/distributions, bars, box plots, pie charts, time series
- Trees, graphs, etc.
  - Nodes and edges
- Geospatial
  - Maps, areas, etc.
- Fields, wordclouds, bubbles, etc.

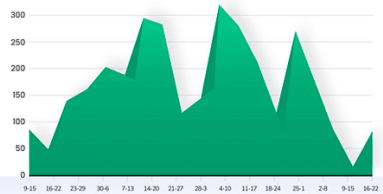
Choice depends on:

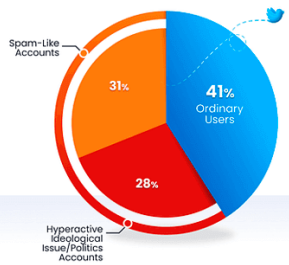
- Nature of data
- Nature of insights a person wants to derive

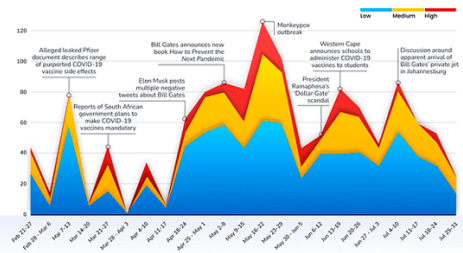

COVID-19
Vaccines
'Dollar Gate' Scandal
29%
28%
17%
7% - Vaccines
3% - Depopulation
3% - National Health Insurance
3% - LGBT
2% - Climate
2% - Mosquito Research
2% - Bill Gates' Private Jet
2% - Synthetic Meat
2% - World Health Organization

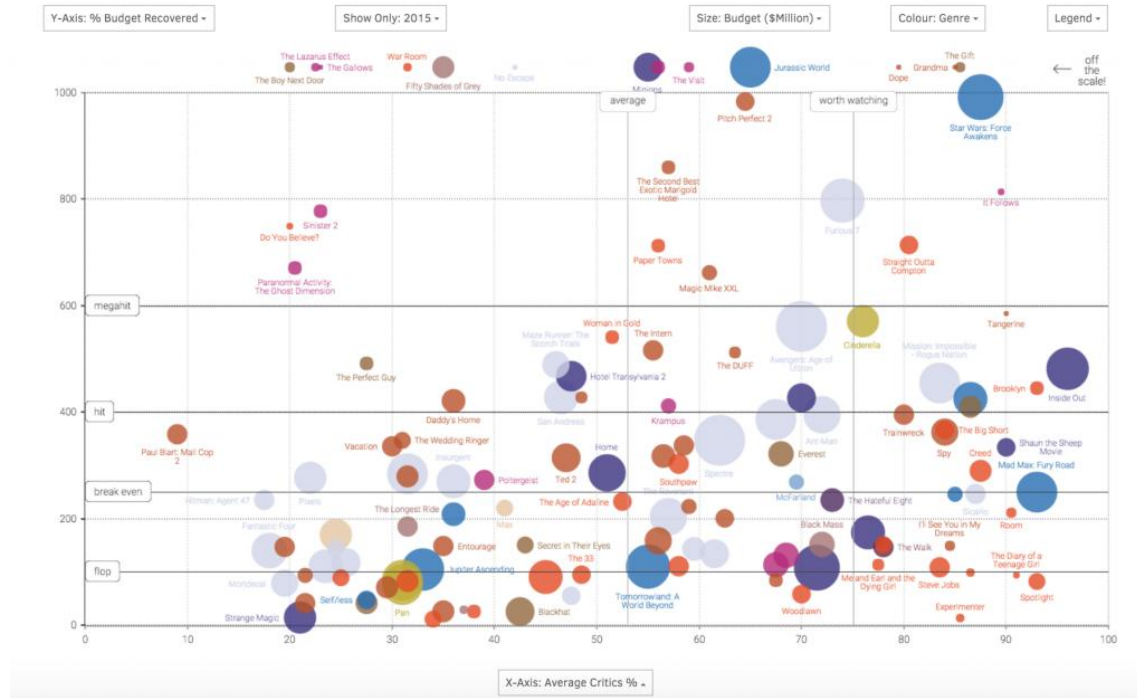

Pollution/Waste
Conservation
Green Technology
Natural Disaster
Government Leadership/ Policy
22%
Climate Cooperation
17%
Energy
11%
General
10%
Food & Water Security
9%
9%
8%
8%
6%

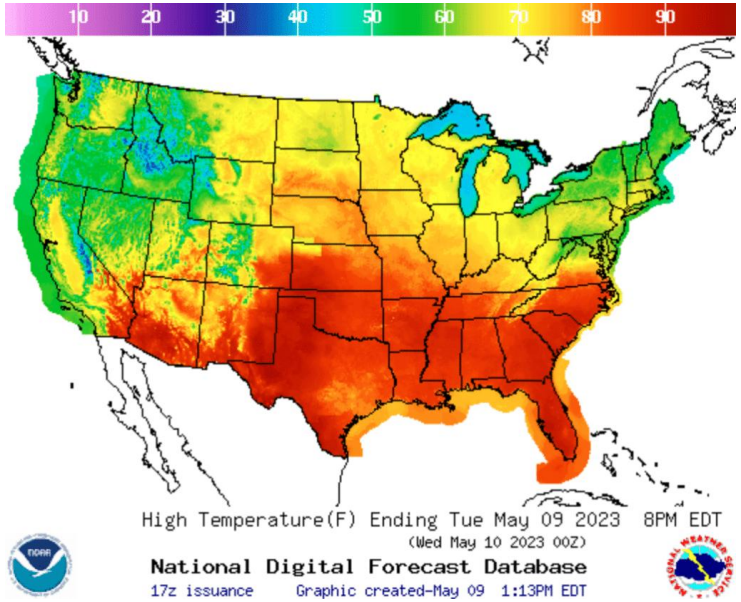

Local Economy
Food / Agriculture
Conspiracy
Military / Security in Africa
Energy / Fuel
General / Other
Sport & Entertainment
Russia Domestic
Bilateral Relations
Media / Propaganda
Russia Leadership
Russia-Ukraine War
Foreign Relations
BREAKDOWN OF THEMES ACROSS ALL COUNTRIES
38%
25%
11%

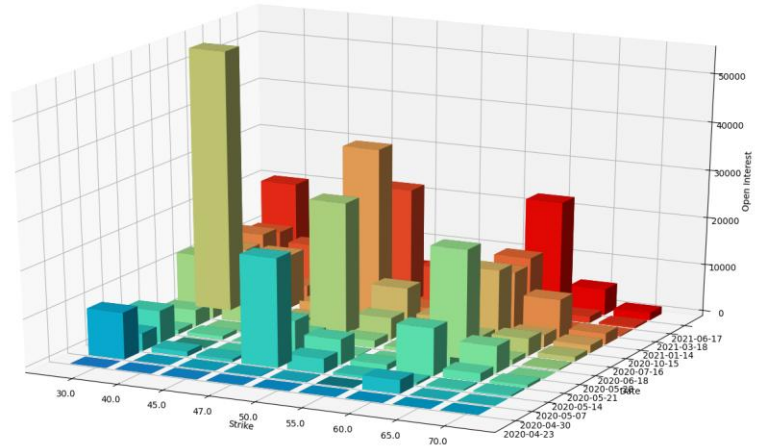

300
250
200
150
100
50
0
9-15
16-22
23-29
30-6
7-13
14-20
21-27
28-3
4-10
11-17
18-24
25-1
2-8
9-15
16-22
April
May
Jun
Jul


Spam-Like Accounts
31%
41%
Ordinary Users
28%
Hyperactive Ideological Issue/Politics Accounts


Low
Medium
High
Alleged leaked Pfizer document describes range of purported COVID-19 vaccine side effects
Reports of South African government plans to make COVID-19 vaccines mandatory
Elon Musk posts multiple negative tweets about Bill Gates
Bill Gates announces new book How to Prevent the Next Pandemic
Monkeypox outbreak
President Ramaphosa's 'Dollar Gate' scandal
Western Cape announces schools to administer COVID-19 vaccines to students
Discussion around apparent arrival of Bill Gates' private jet in Johannesburg


Y-Axis: % Budget Recovered
Show Only: 2015
Size: Budget ($Million)
Colour: Genre
Legend
off the scale!
average
worth watching
megahit
hit
break even
flop
X-Axis: Average Critics %


10
20
30
40
50
60
70
80
90
High Temperature(F) Ending Tue May 09 2023 8PM EDT
(Wed May 10 2023 00Z)
National Digital Forecast Database
17z issuance
Graphic created-May 09 1:13PM EDT


Open Interest
Strike
Date

wavelengths
environment
given
original
modeling
similar
optimal
scores
time
authors
developed
BioCompute
platform
approach
samples
application
dataset
standard
proposed
example
regression
intervals
Dockerfile
LogD
version
variable
study
related
code
high
prediction
review
method
network
system
shows
vector
provided
molecular
test
available
project
community
query
run
create
Shiny
ADR
cross
number
OHPL
similarity
functions
calibration
drugs
methods
plot
workflow
group
PLS
containers
development
sequence
types
user
interval
ElasticNet
local
model
specific
partial
software
coefficient
allows
generate
validation
step
models
new
groups
tools
CGC
support
portal
computational
provides
iPLS
space
information
data
value
analysis
process
access
datasets
performance
different
OSAR
error
adaptive
common
set
large
values
workflows
selected
app
variables
selection
generated
web
submission
package
lasso
squares
users
drug
training
multiple
work
packages
SVM
PDXNet
structure
Docker
descriptors
research
function
parameter
testing
details
tool
provide
machine
algorithm
protein
features
including
bioinformatics
important
section
following
container
dimensional
environments
domain
evaluation
include
fields

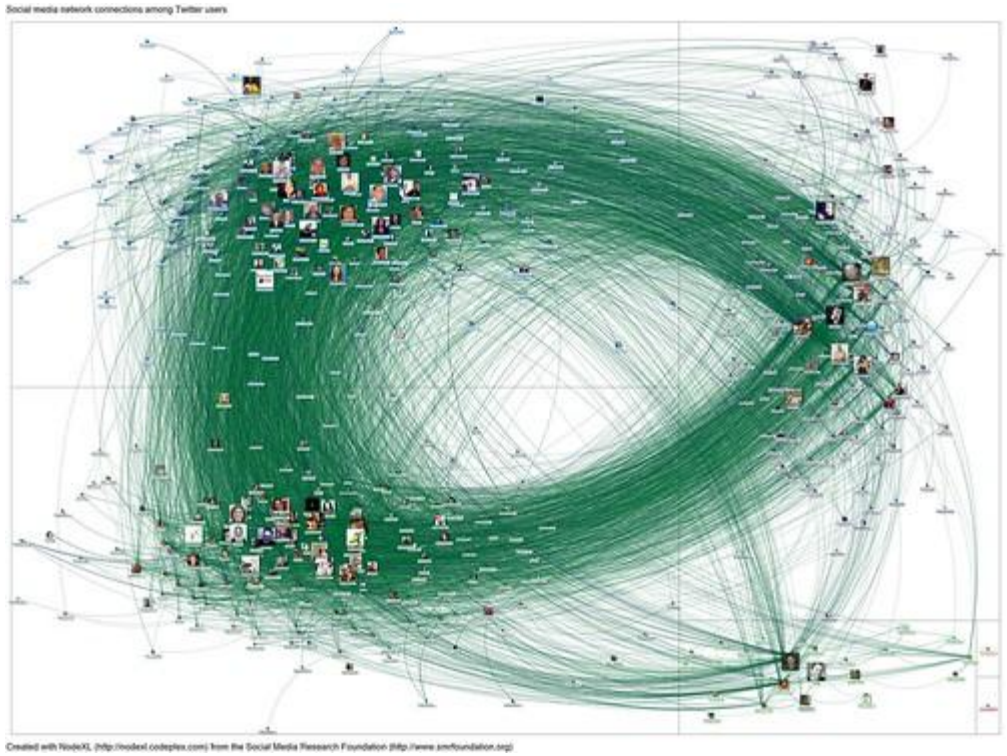
How do you go about designing visualizations

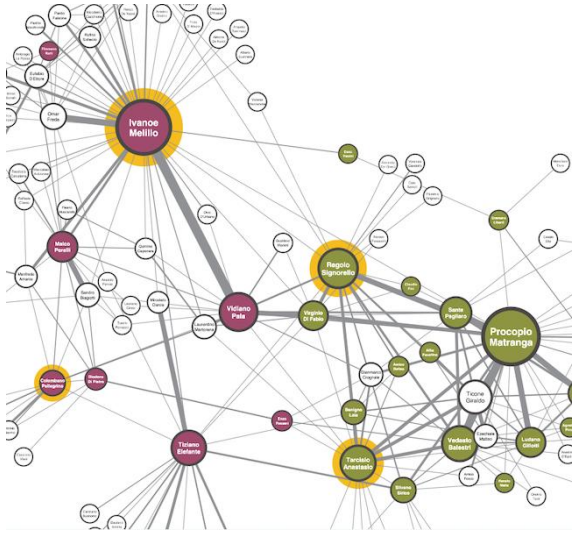
- Start with data
- What do people want out of the data?
  - And who are the people?

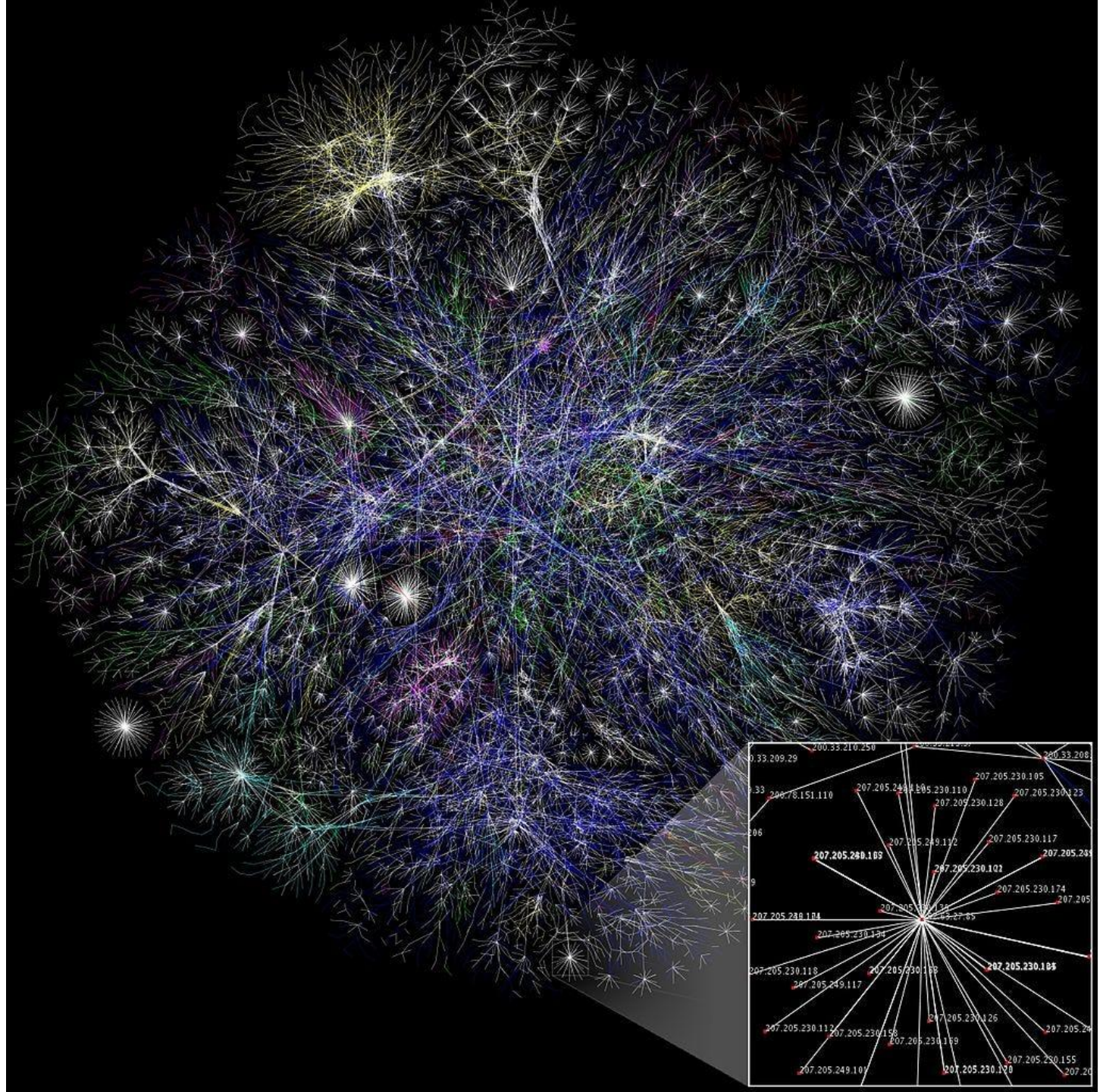
- E.g., Train locations and movements
  - What a traveller wants vs. traffic management wants

- Pick visualizations that communicate that insight to users
- Set of techniques to break down complexity

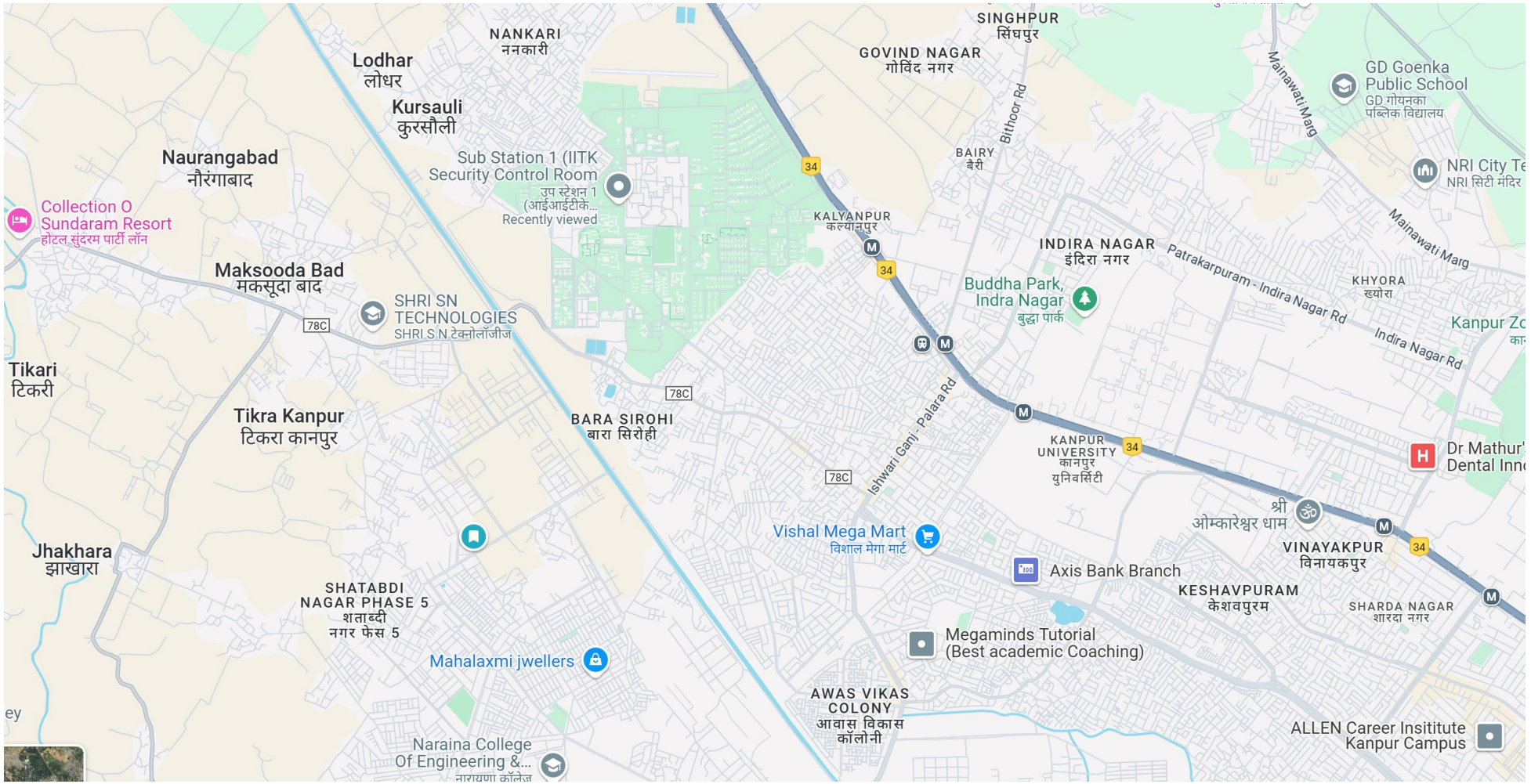
Some strategies: Micro vs. Macro level views



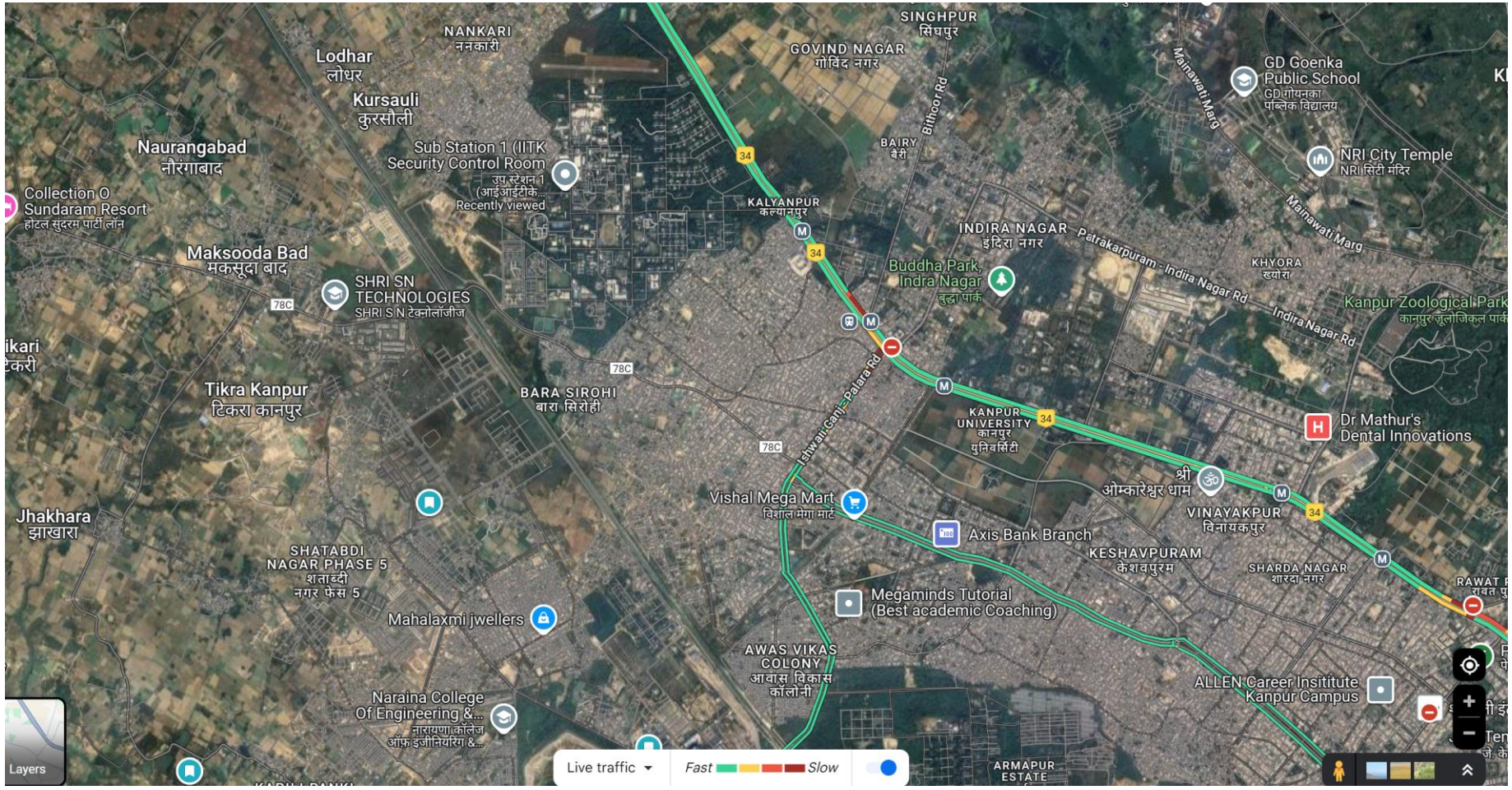



200.33.210.250
207.205.230.105
207.205.230.110
207.205.230.123
200.78.151.110
207.205.230.128
207.205.230.117
207.205.249.112
207.205.230.174
207.205.230.118
207.205.249.117
207.205.230.126
207.205.230.112
207.205.230.158
207.205.230.169
207.205.230.155
207.205.249.107

Some strategies: Layering



Some strategies: Layering



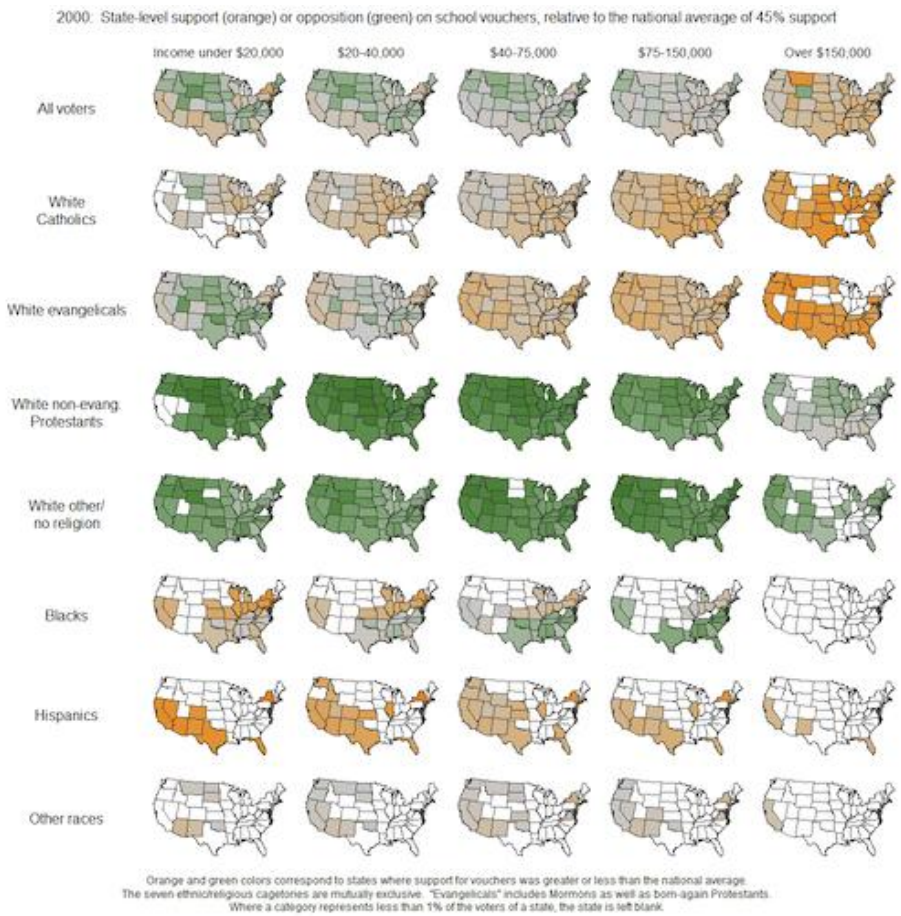
Some strategies: Separation

- Show one route at a time
- Show only between one source and destination
- All routes from/to a station
- Only no toll roads
- Etc.

More separation

- https://www.chartjs.org/docs/latest/samples/bar/border-radius.html

Small multiples

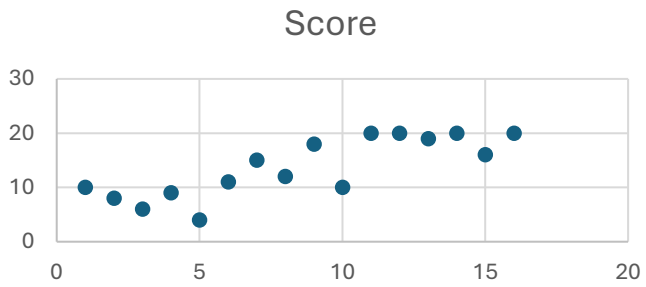


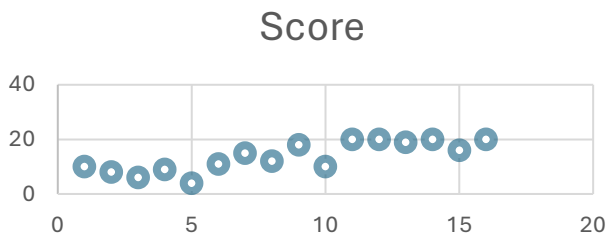
Summary of complexity taming strategies

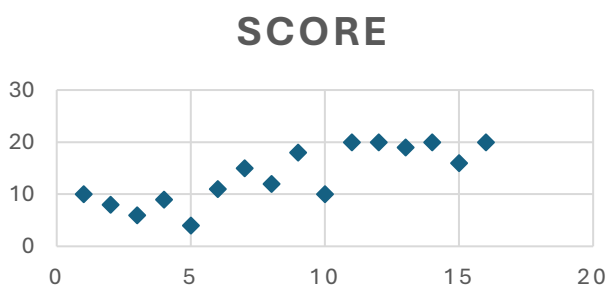
- Micro – macro views
- Layering
- Separation
- Small multiples

Designing visualizations: Primitives

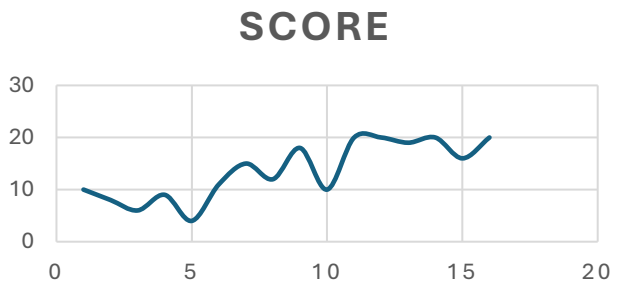
Points

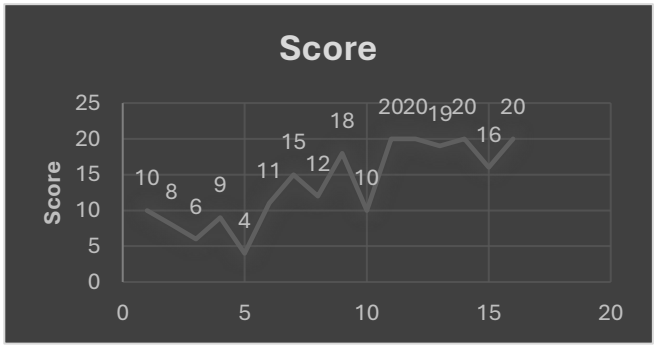




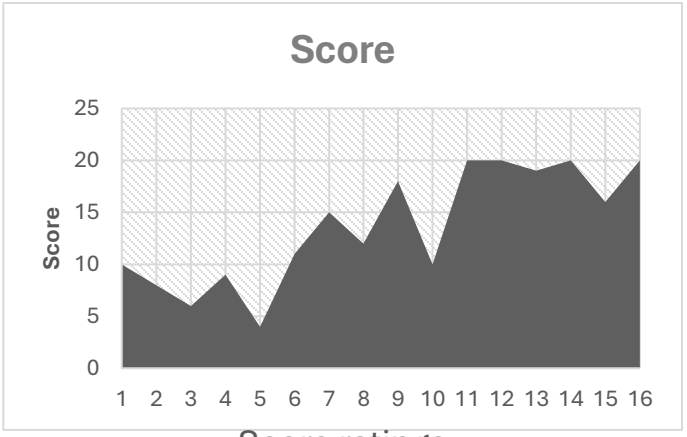


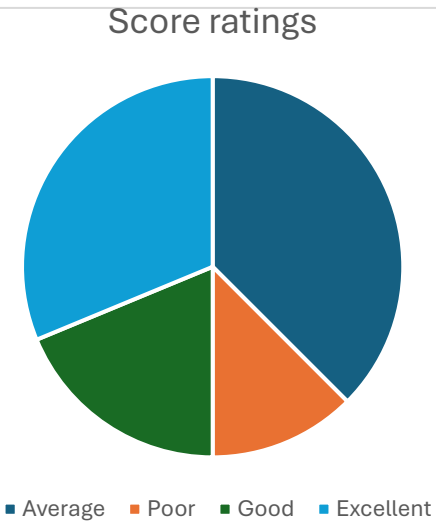
Lines





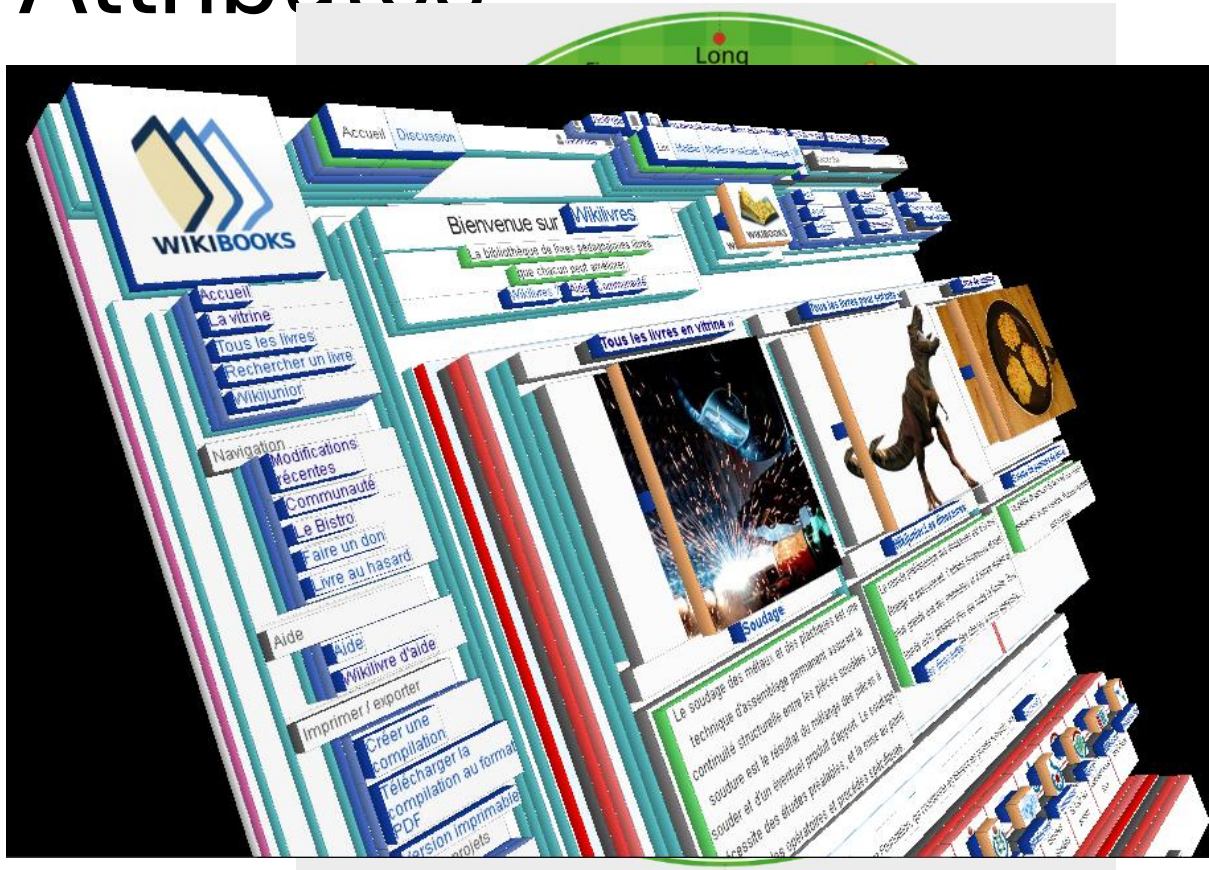
Area





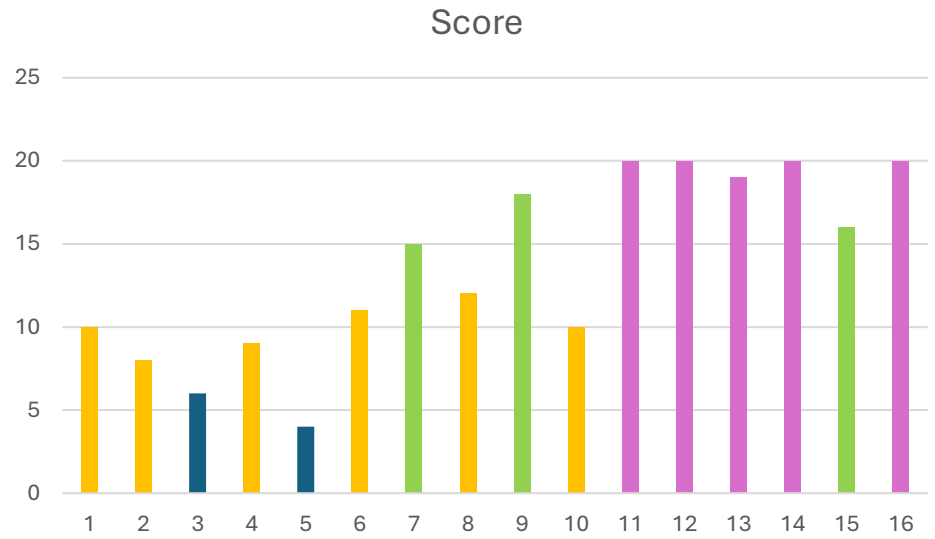
Designing visualizations: Attributes

- Shape
- Size
- Position
- Color
- Thickness/weight
- Tilt
- Area/Volume
- Direction
- Additionally-labels, axes, titles



Options you see in “Insert chart” are some variations of these!

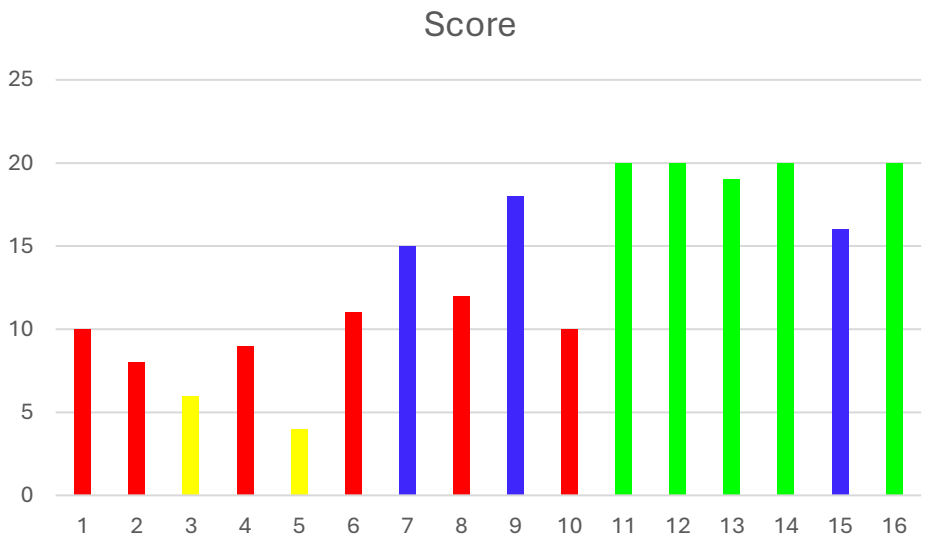
Rules of thumb

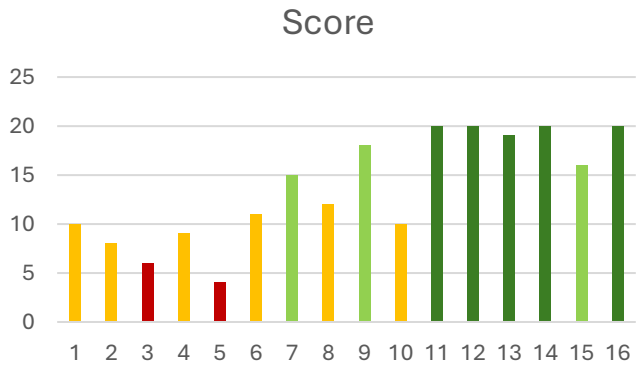


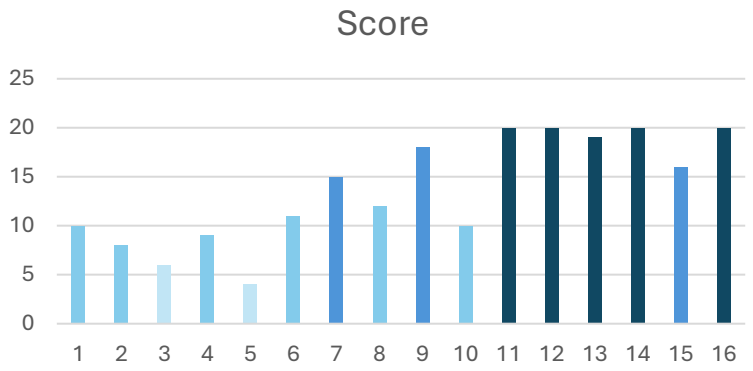
- One dimension, one notation
  - If you use colors to represent score…
    - Use only colors for scores
    - Use colors only for scores, and nothing else.
  - If you use height => use only height.
  - Using color and height is okay…
    - PROVIDED you don’t need colors / heights for anything else.
  - Same with shape, thickness, etc.
- Also allows for comparison
  - E.g., if you use color category, and dots/dashes for gender, you could compare across categories and gender [same color different patterns, same pattern different color comparisons].

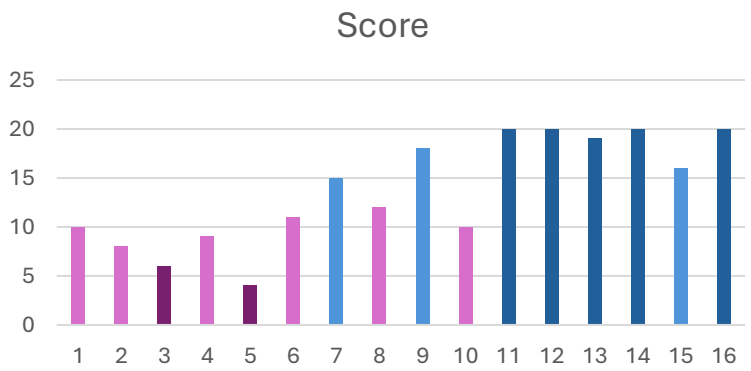
Rules of thumb

- No bright colors
  - Eyes are super sensitive to colors









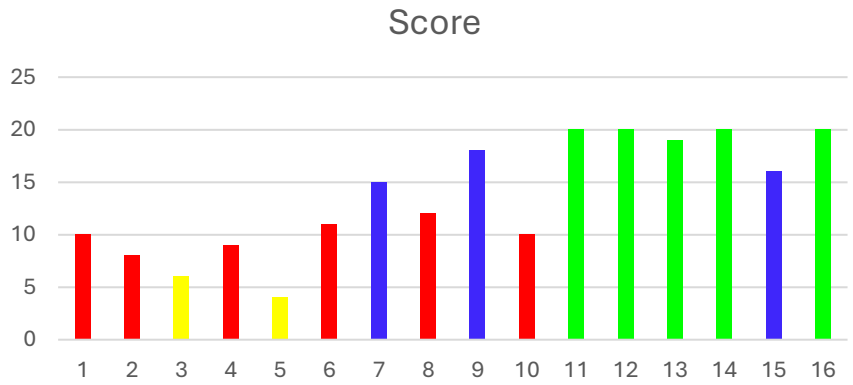
Rules of thumb

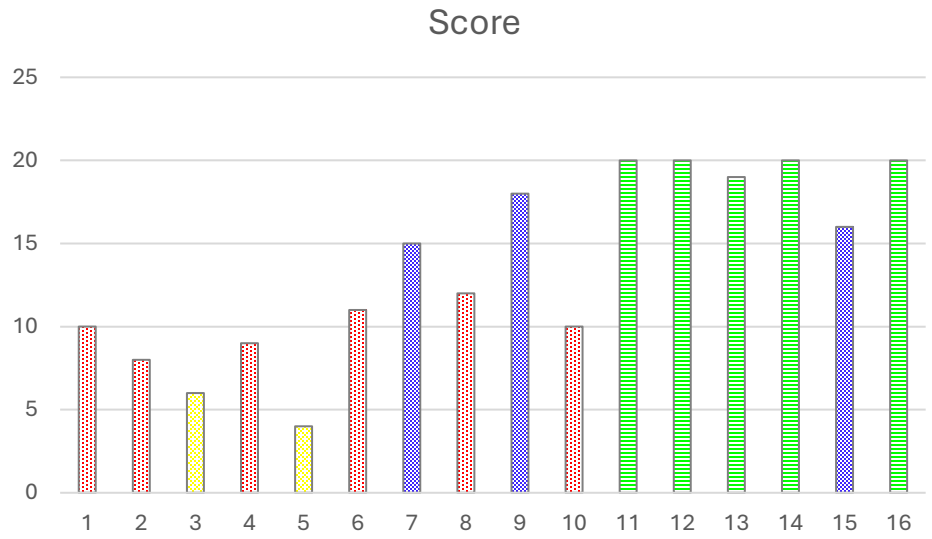
- When not enough colors, use bright colors judiciously:
  - Smaller areas
  - Draw attention
  - Reduce width/size

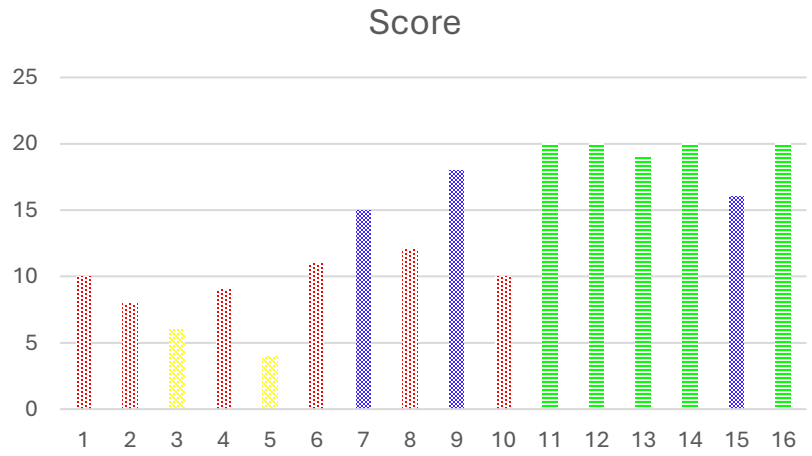
- Never put them next to white or black!
- More on color theory later!!

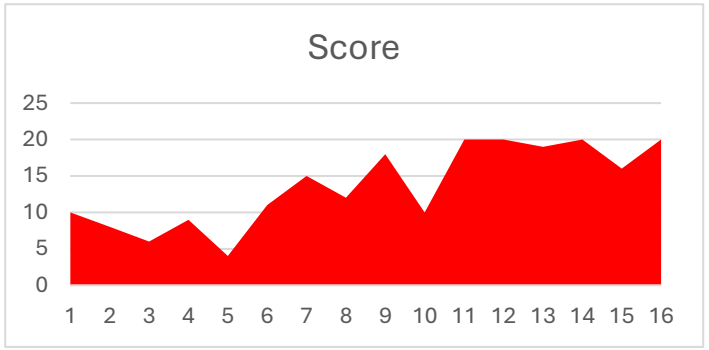
Rule of thumb

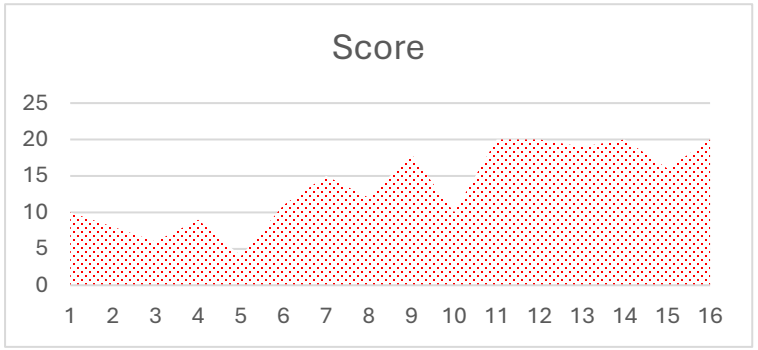
- Use shades, gradients to lessen/increase effects of colors

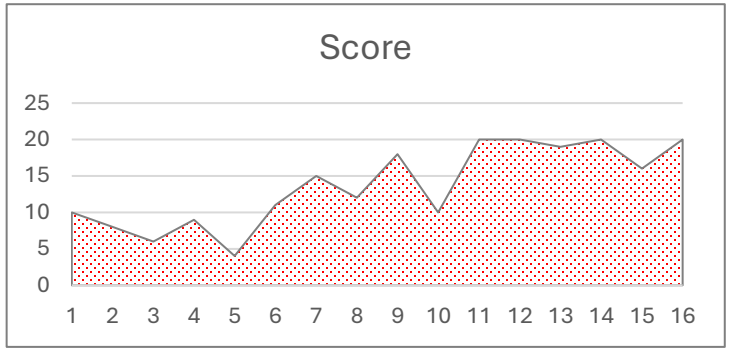


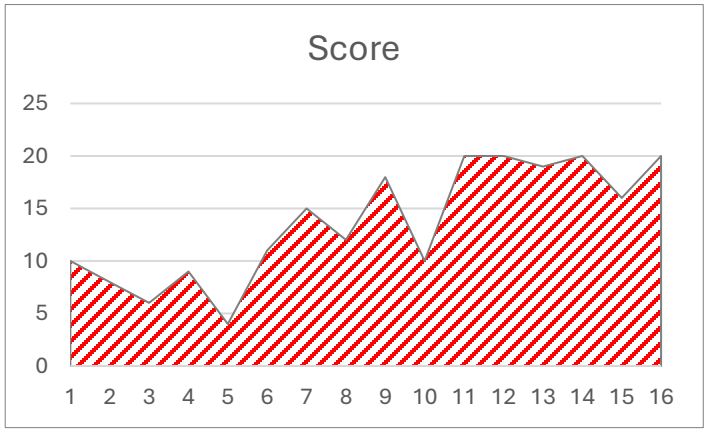


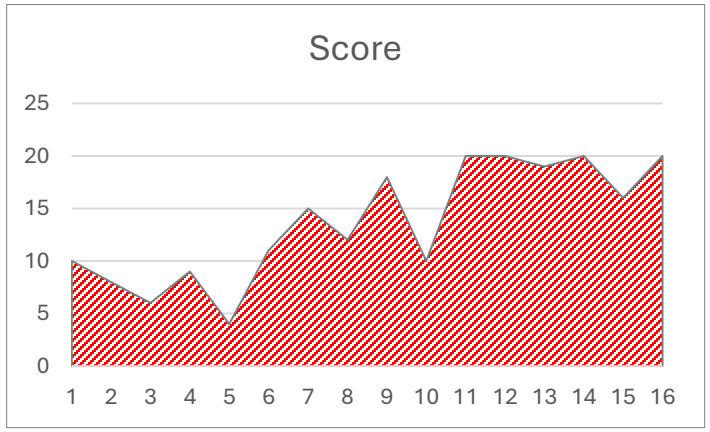


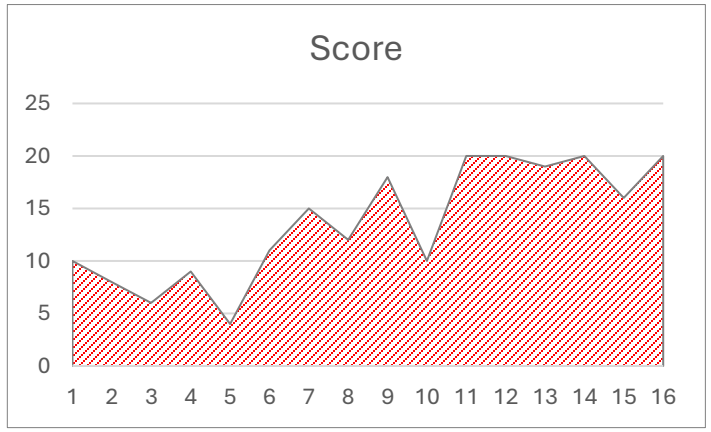

Score
25
20
15
10
5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16


Score
25
20
15
10
5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16


Score
25
20
15
10
5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

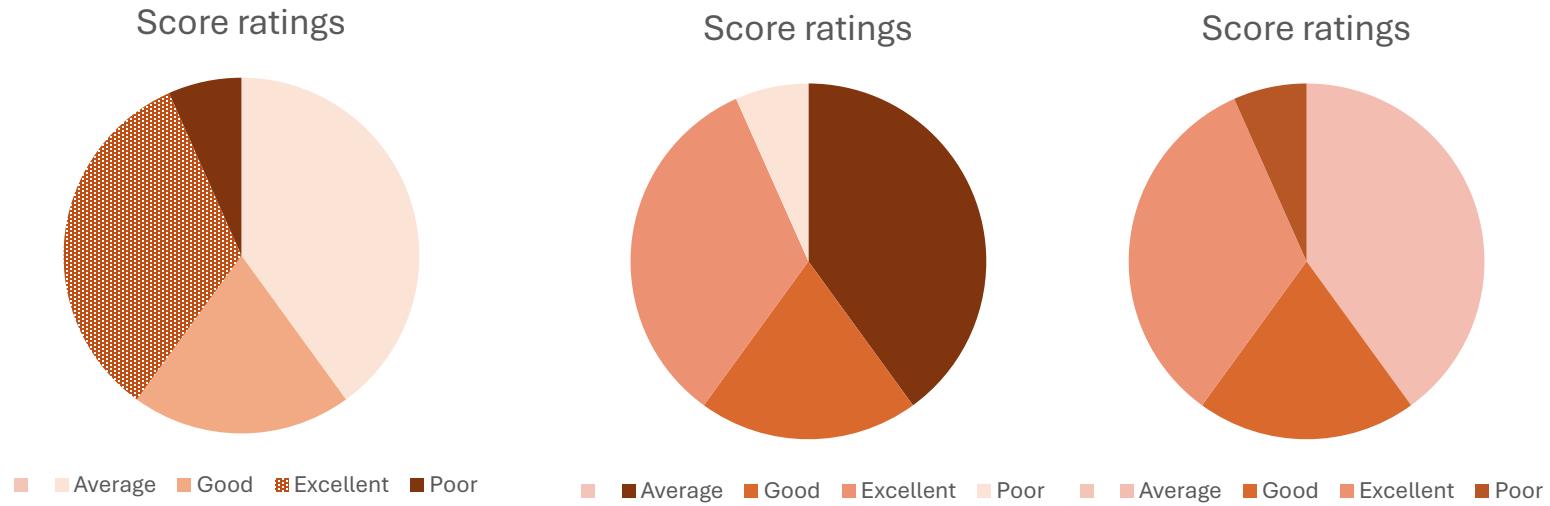

Score
25
20
15
10
5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16


Score
25
20
15
10
5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16


Score
25
20
15
10
5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

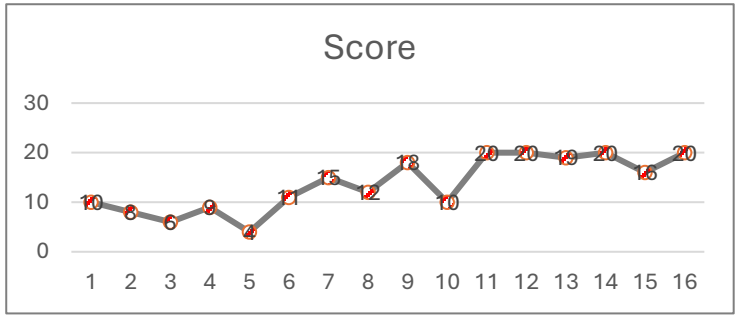
Rule of thumb

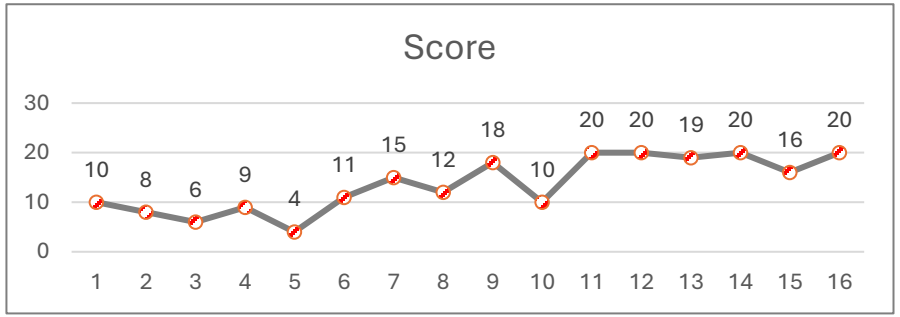
- Think of grayscale (for B/W printing) and contrasts always!
  - Which is better?

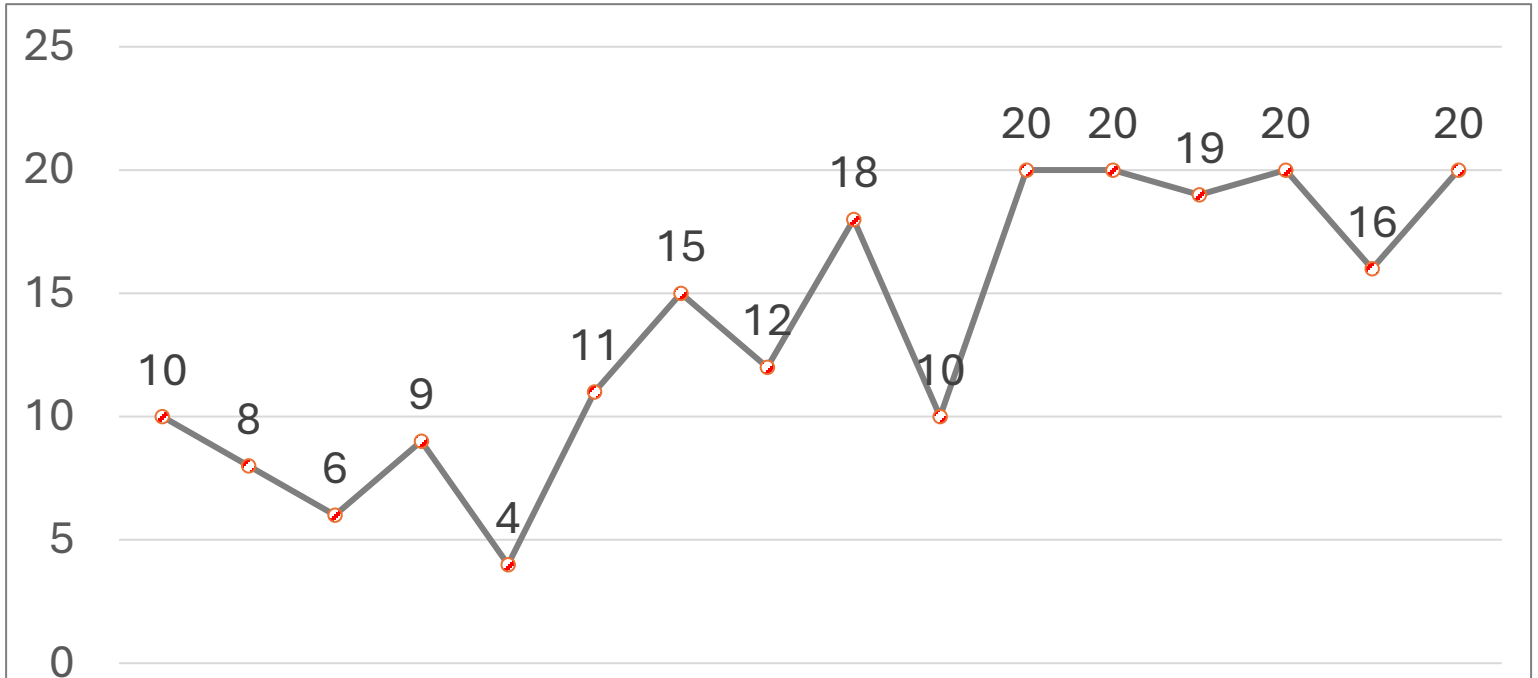


Rule of thumb

- Avoid overlaps
- Keep labels readable
- Reading size: 10-12 on paper/documents, 18+ for presentations

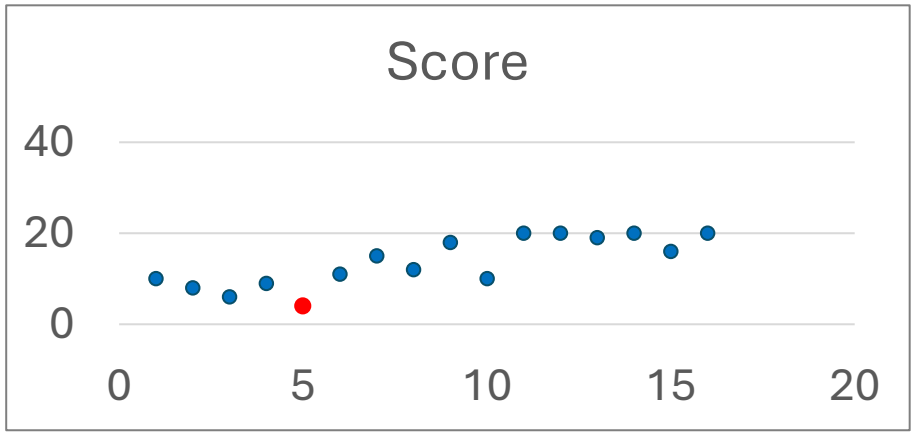


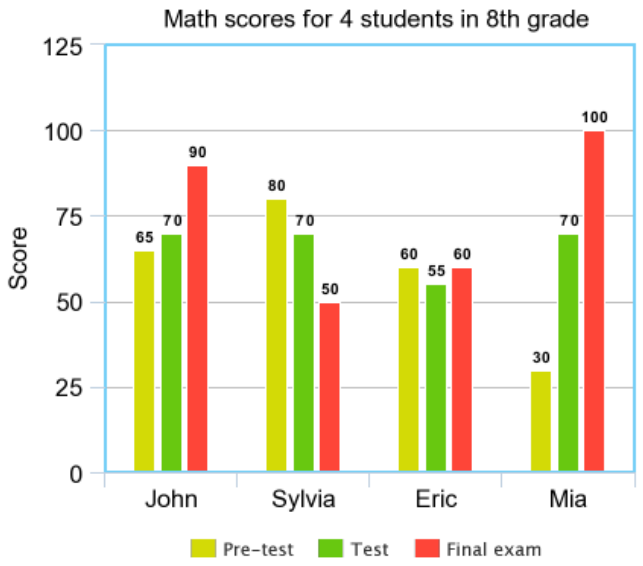


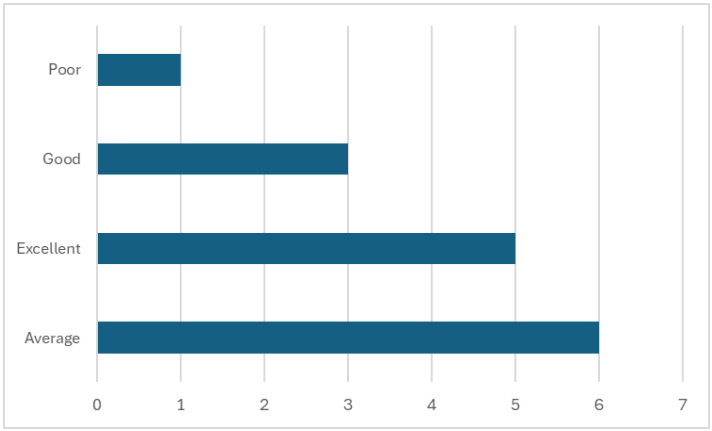


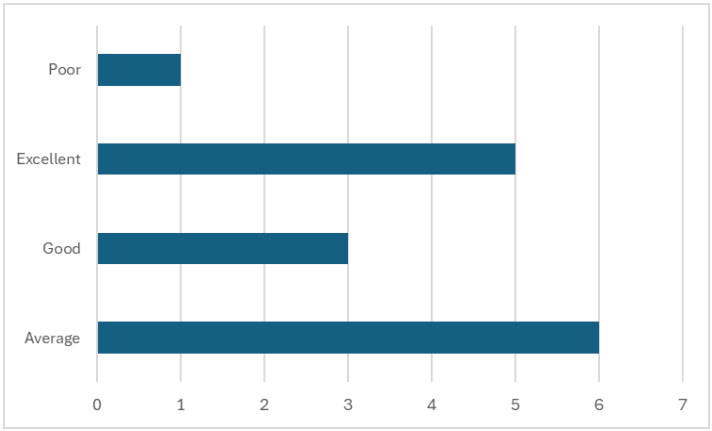
Rule of thumb

- Use groupings, sortings, popouts to your advantage

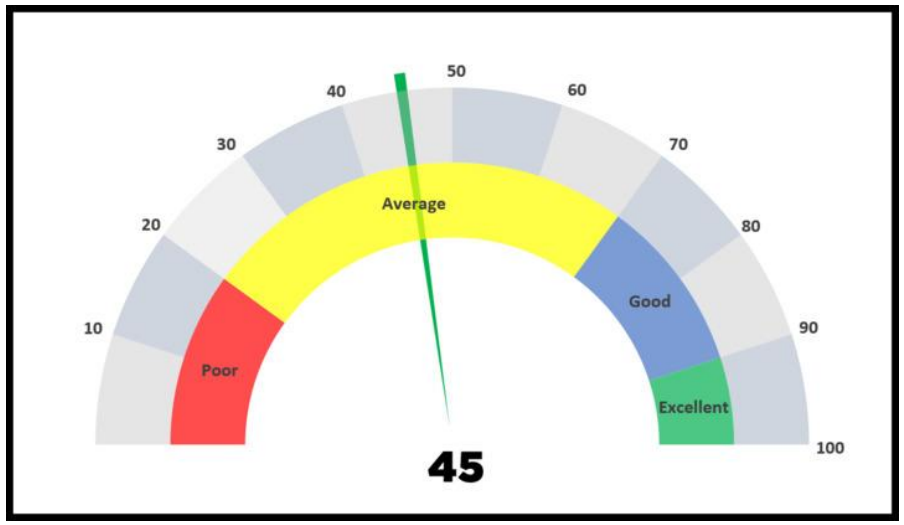


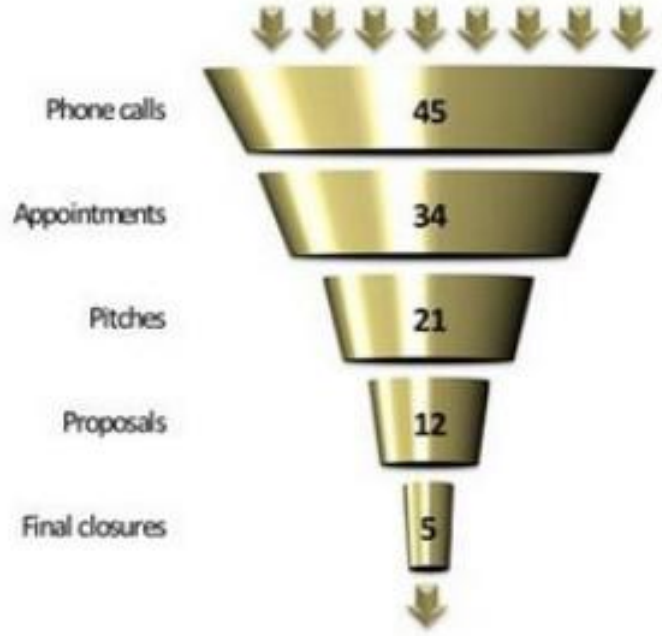


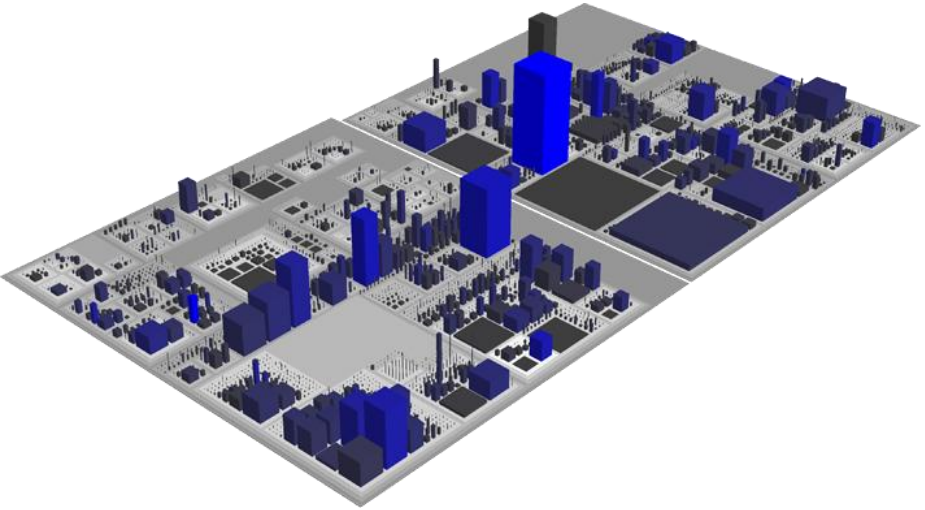


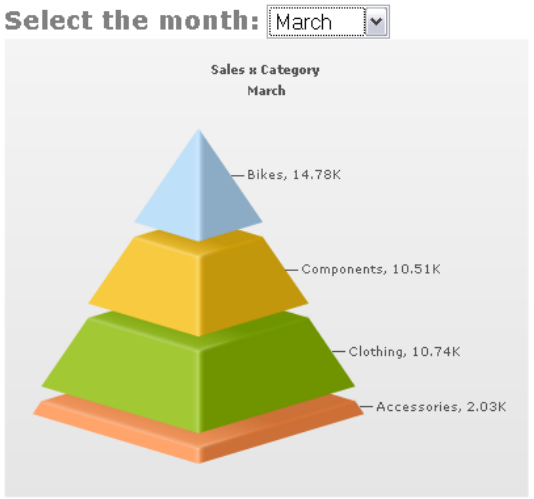


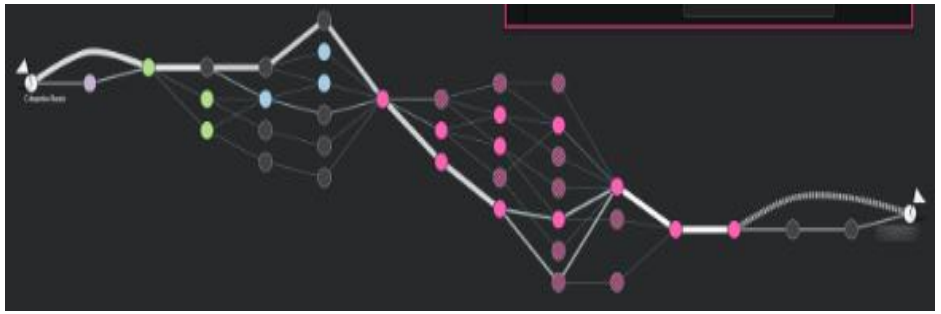
Look for natural mapping / familiar metaphors



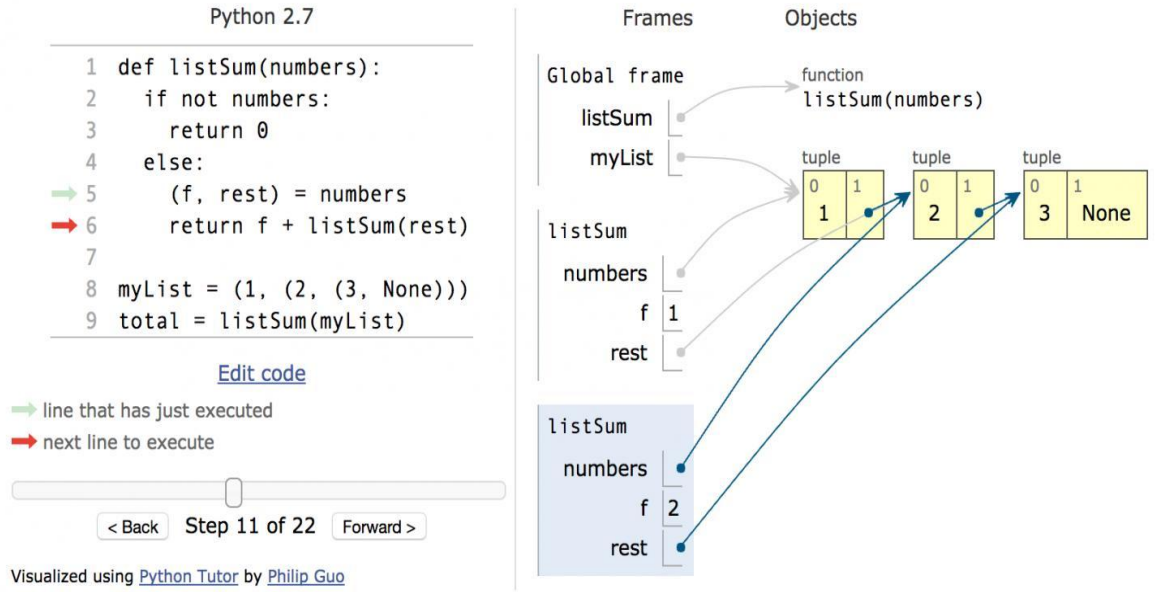




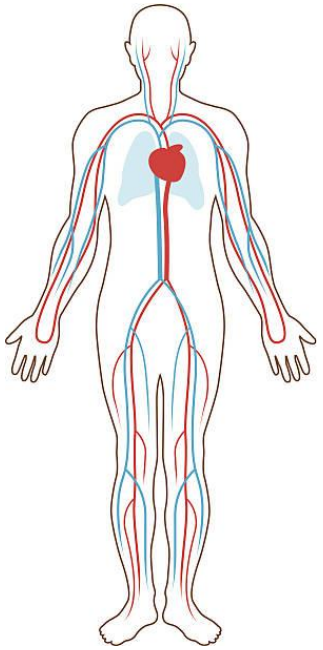


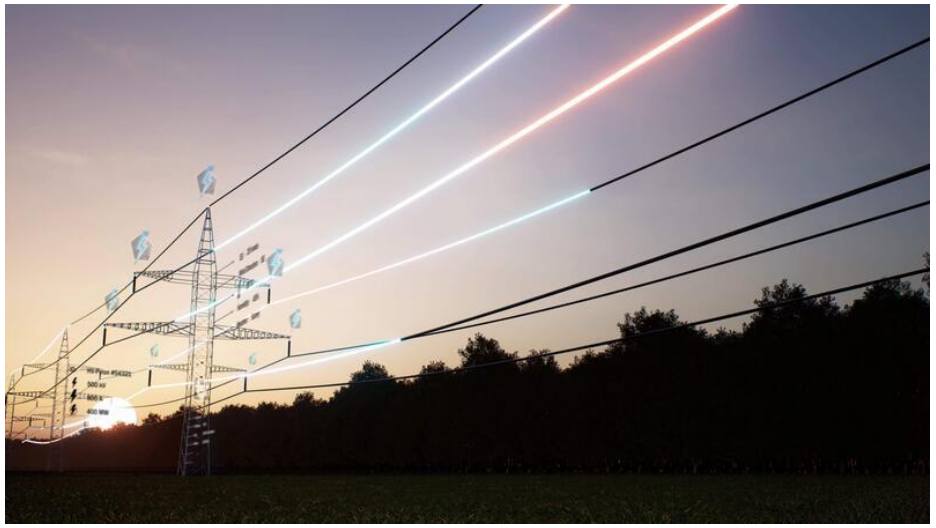
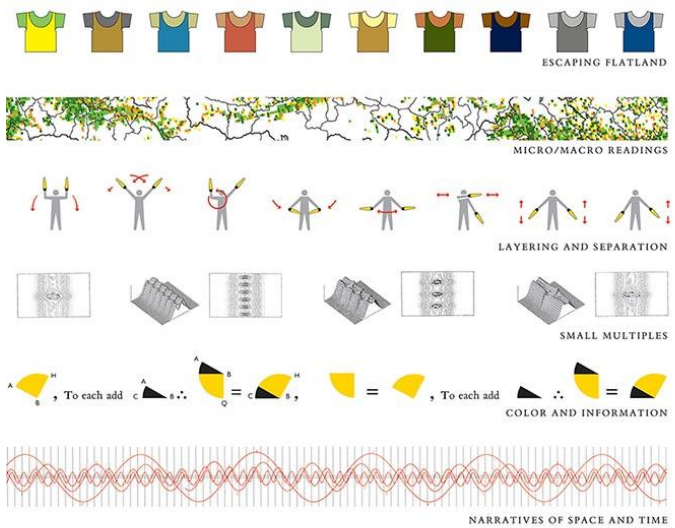
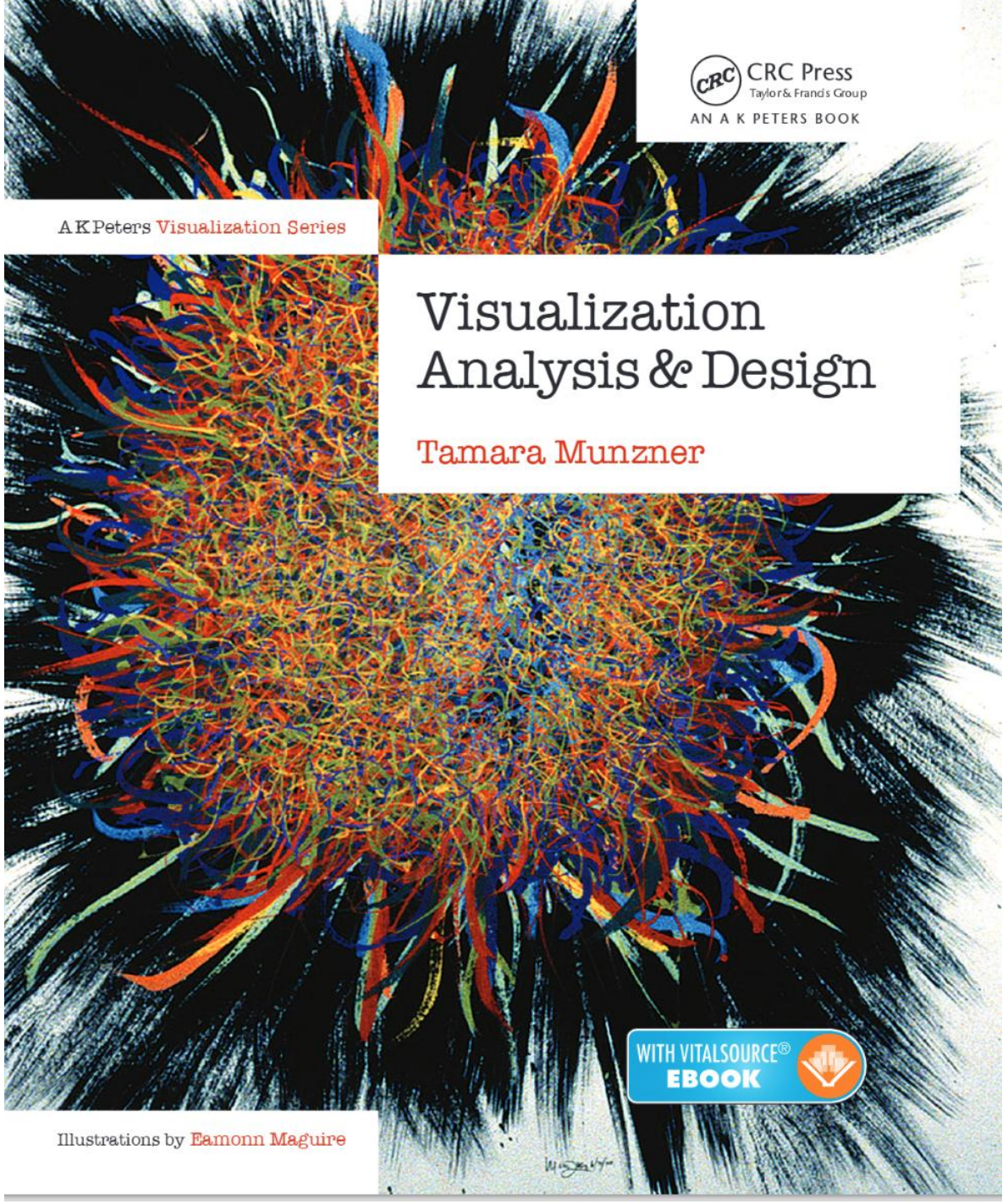


Visualizations to reflect reality



Visualizations to reflect reality





Additional Reading