Prof. Gerald is studying the effect of the chemical compound commonly known as De Vries' Extract on pupil radius on different closely-related species.

He collects four samples of compound intensity and pupil radius across four species: Bruxa, Striga, Basilisk, and Kikimore.

He suspects a positive linear relationship between intensity and radius: wider pupils as more of the compound is ingested.

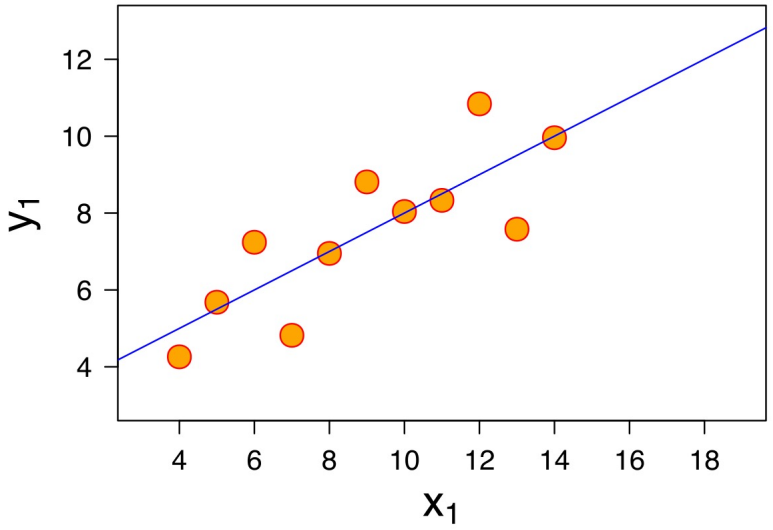
Help him analyze the properties of each dimension and the relationships between them.

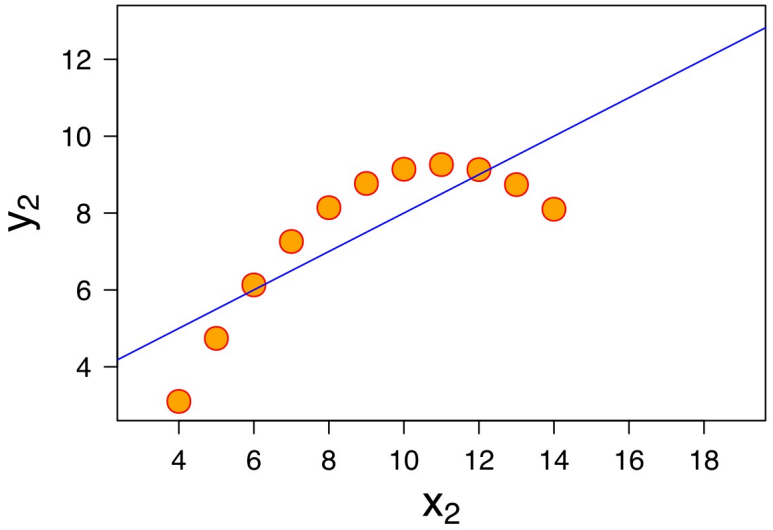
Anscombe's quartet

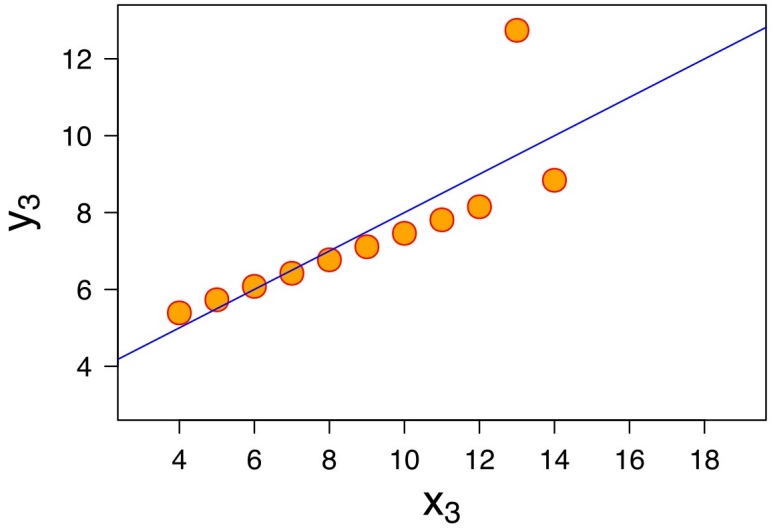
I		II		III		IV	
x	*y*	*x*	*y*	*x*	*y*	*x*	*y*
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

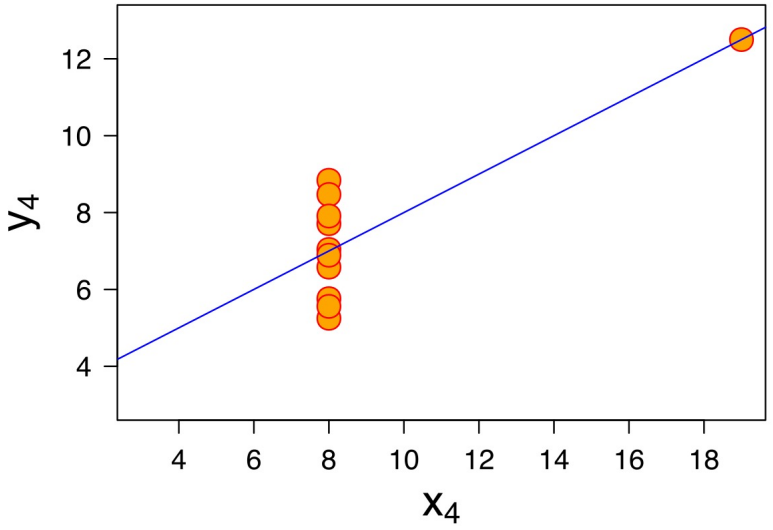
Statistical Property	
Mean of x	9
Sample variance of x : s^2	11
Mean of y	7.50
Sample variance of y : s^2	4.125
Correlation between x and y	0.816
Linear regression line	$y = 3.00 + 0.500x$
Coefficient of determination of the linear regression R^2	0.67

Are these datasets the same?









The importance of seeing your data

Same Stats; Different Graphs

https://www.autodesk.com/research/publications/same-stats-different-graphs

Why do we visualize data?

Answer or discover questions

Make decisions

Contextualize

Expand memory

Aid cognition

Tell a story

Inform

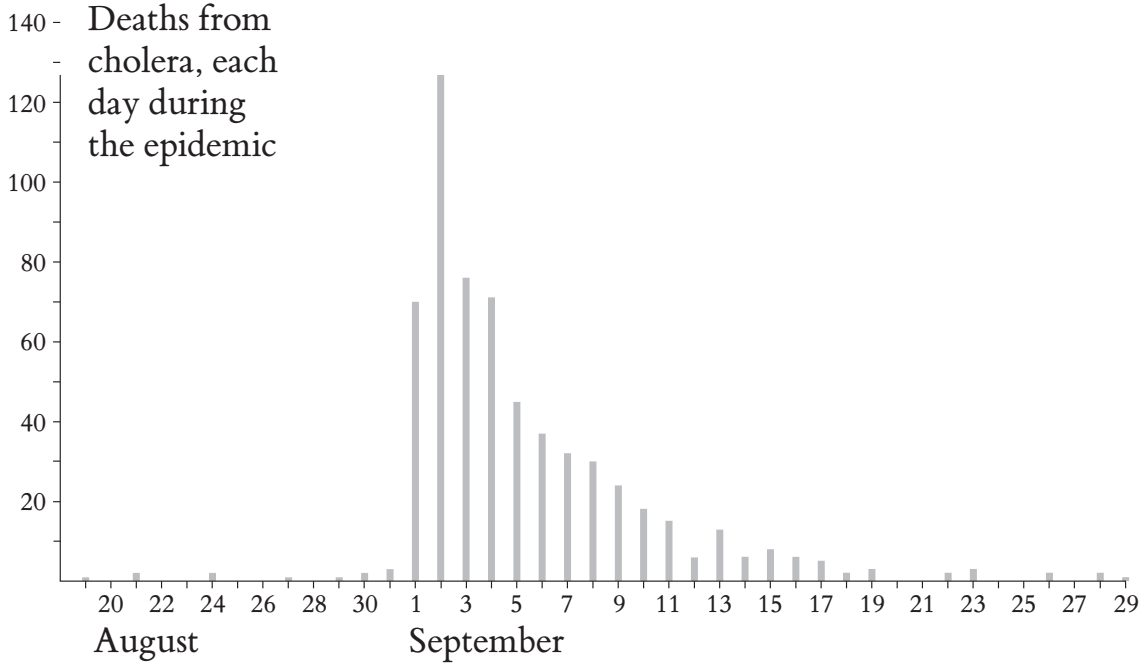
Inspire

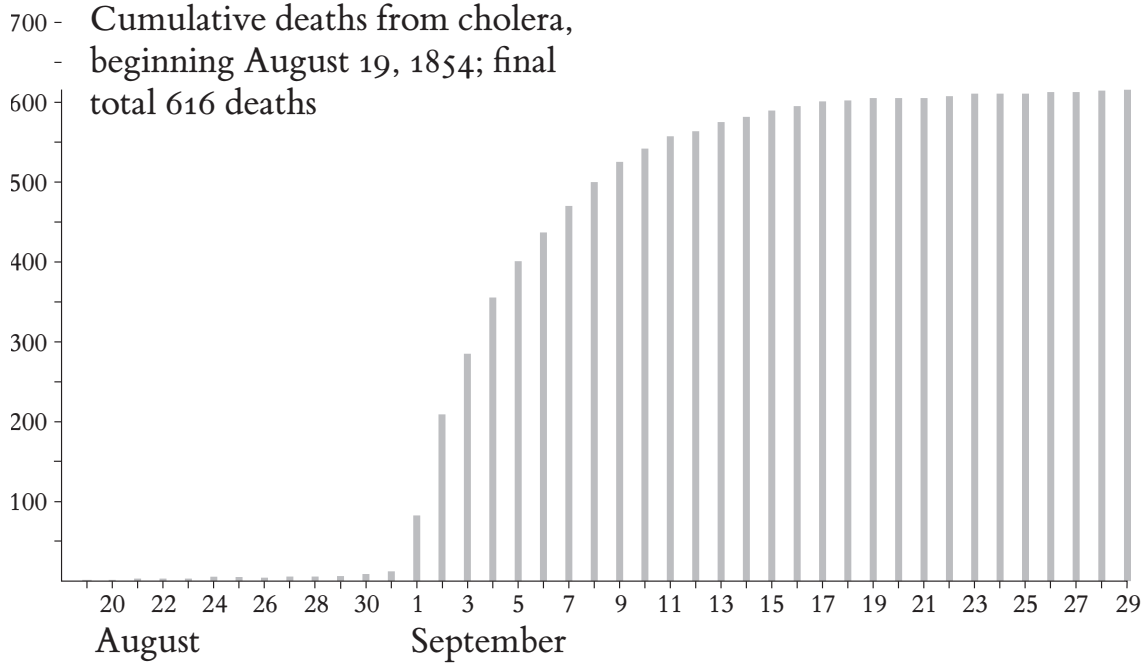
Find patterns

Collaborate

Revise

Find errors

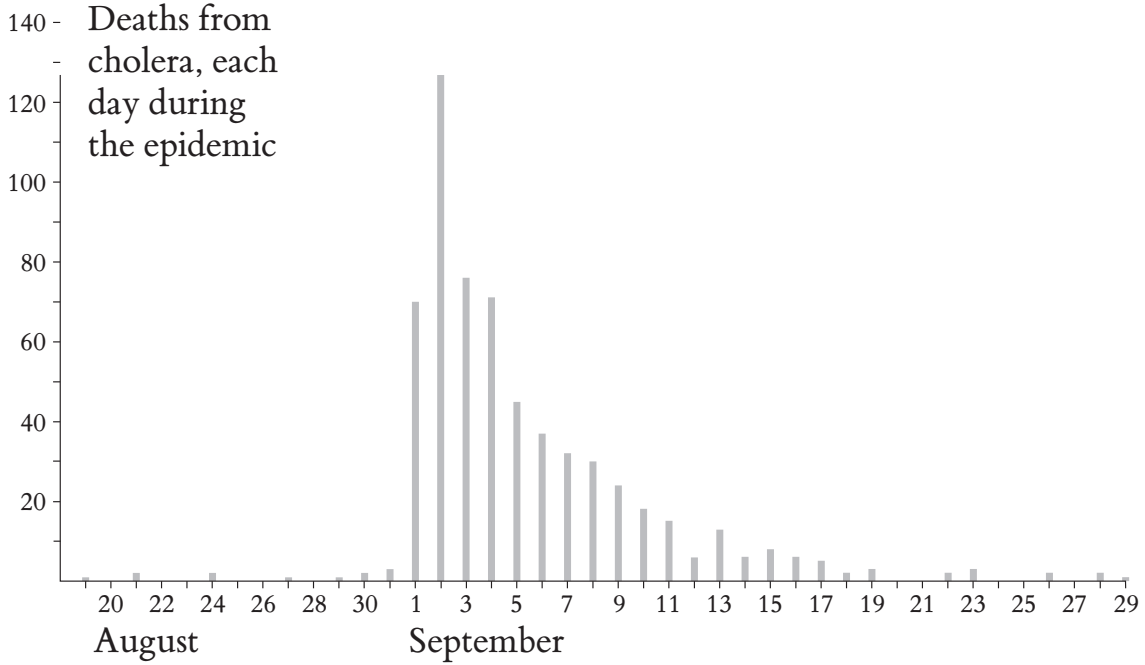


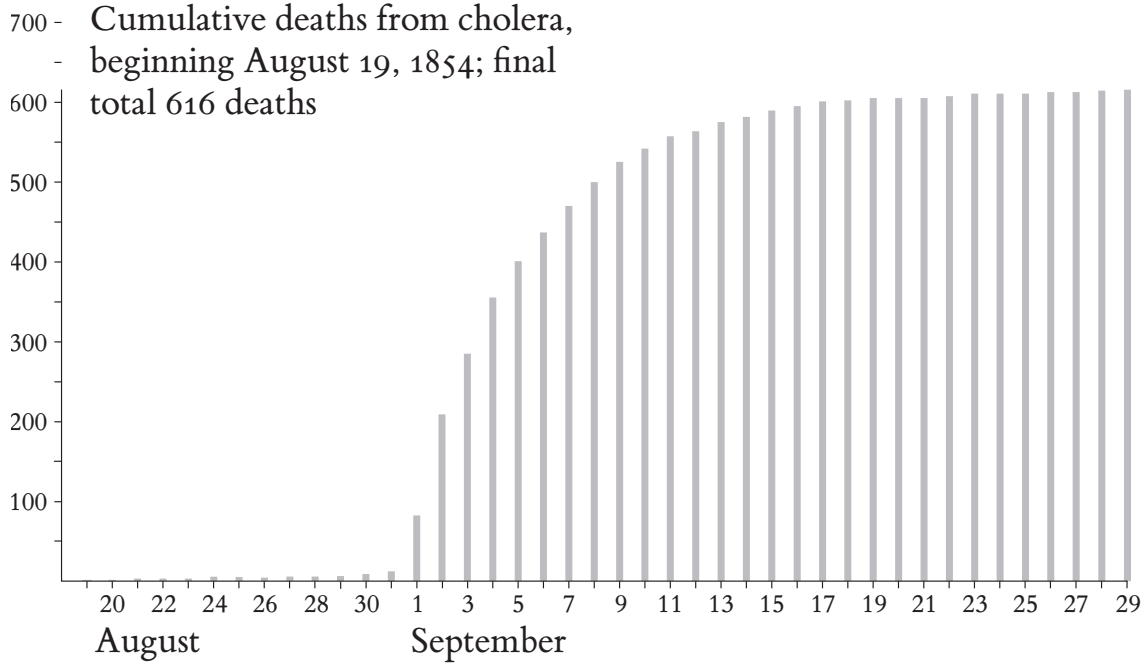


In September 1854, a cholera outbreak resulted in 616 deaths in central London.

The prevailing theory of the time was that disease was airborne.

Support Reasoning





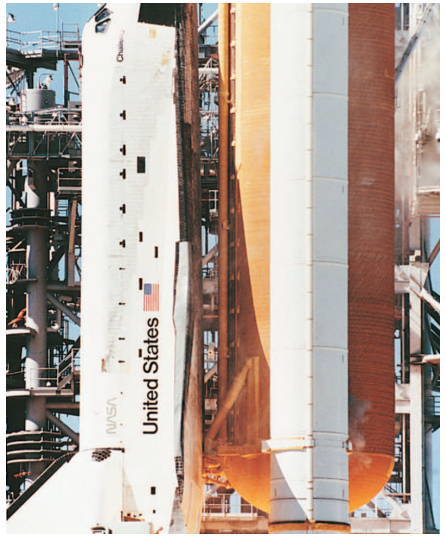
In September 1854, a cholera outbreak resulted in 616 deaths in central London.

The prevailing theory of the time was that disease was airborne.

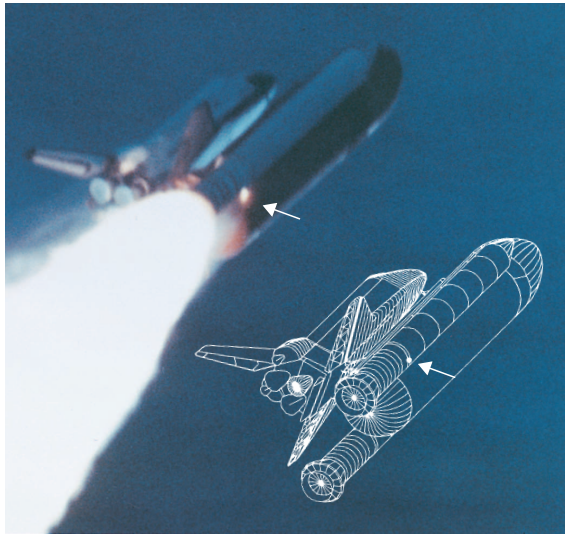
Support Reasoning

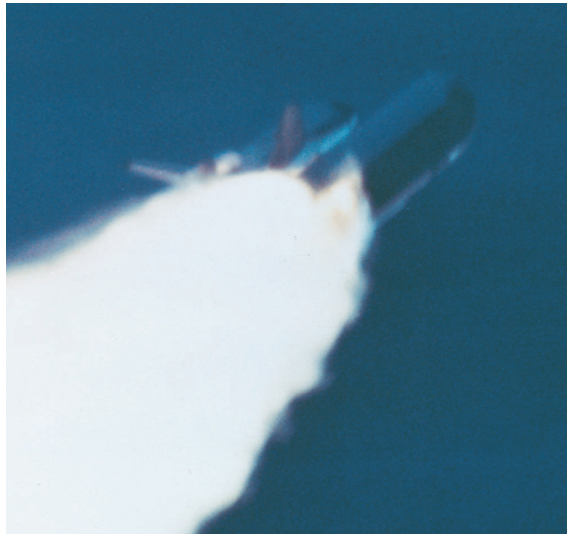
John Snow's scientific detective work aimed to establish causal-effect between dirty water and the outbreak by plotting deaths and showing that they clustered around the broad street pump.

Support Reasoning



Less than 1 second after ignition, a puff of smoke appeared at the aft joint of the right booster, indicating that the O-rings burned through and failed to seal. At this point, all was lost.



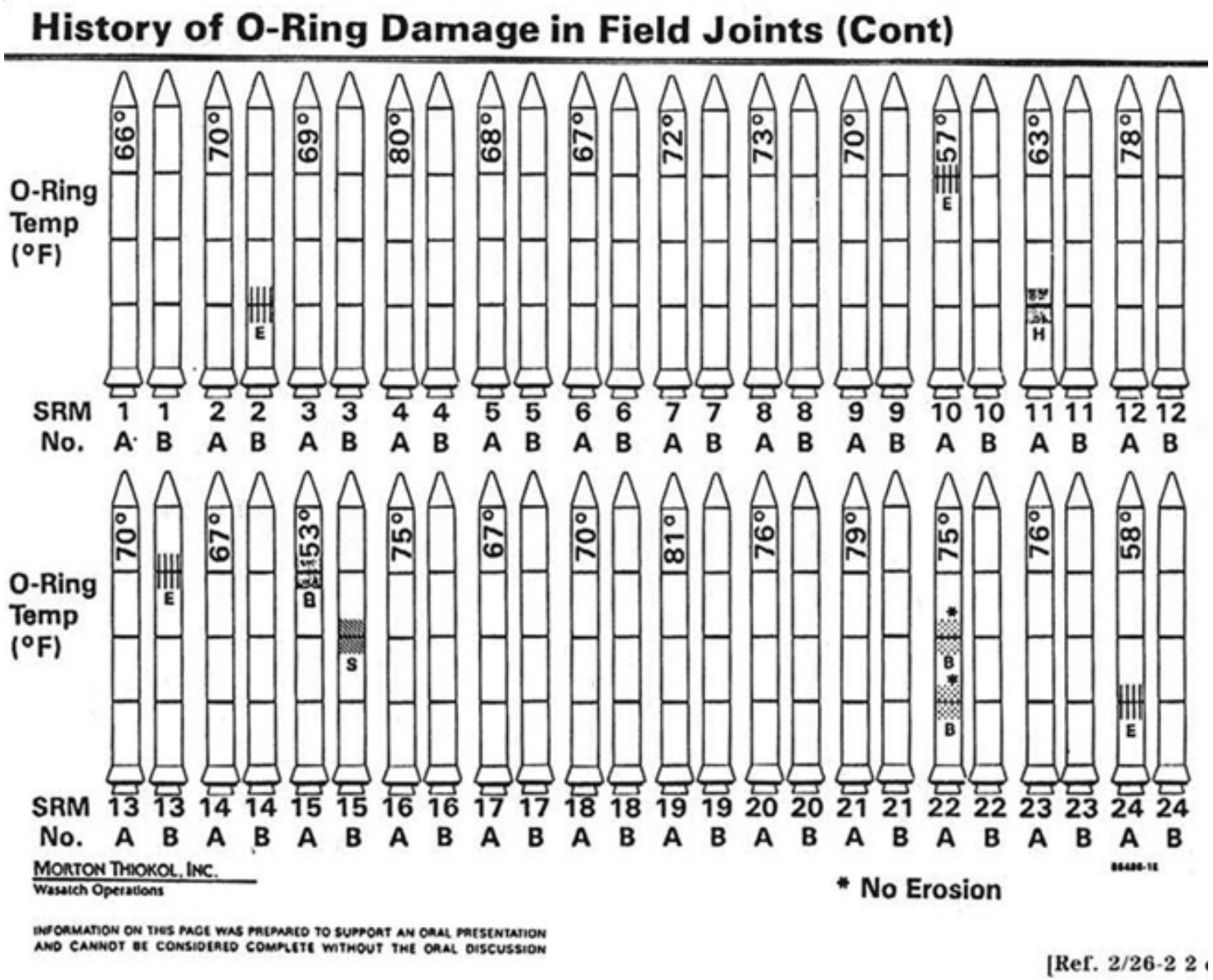


On the launch pad, the leak lasted only about 2 seconds and then apparently was plugged by putty and insulation as the shuttle rose, flying through rather strong cross-winds. Then 58.788 seconds after ignition, when the Challenger was 6 miles up, a flicker of flame emerged from the leaky joint. Within seconds, the flame grew and engulfed the fuel tank (containing liquid hydrogen and liquid oxygen). That tank ruptured and exploded, destroying the shuttle.

As the shuttle exploded and broke up at approximately 73 seconds after launch, the two booster rockets crisscrossed and continued flying wildly. The right booster, identifiable by its failure plume, is now to the left of its non-defective counterpart.

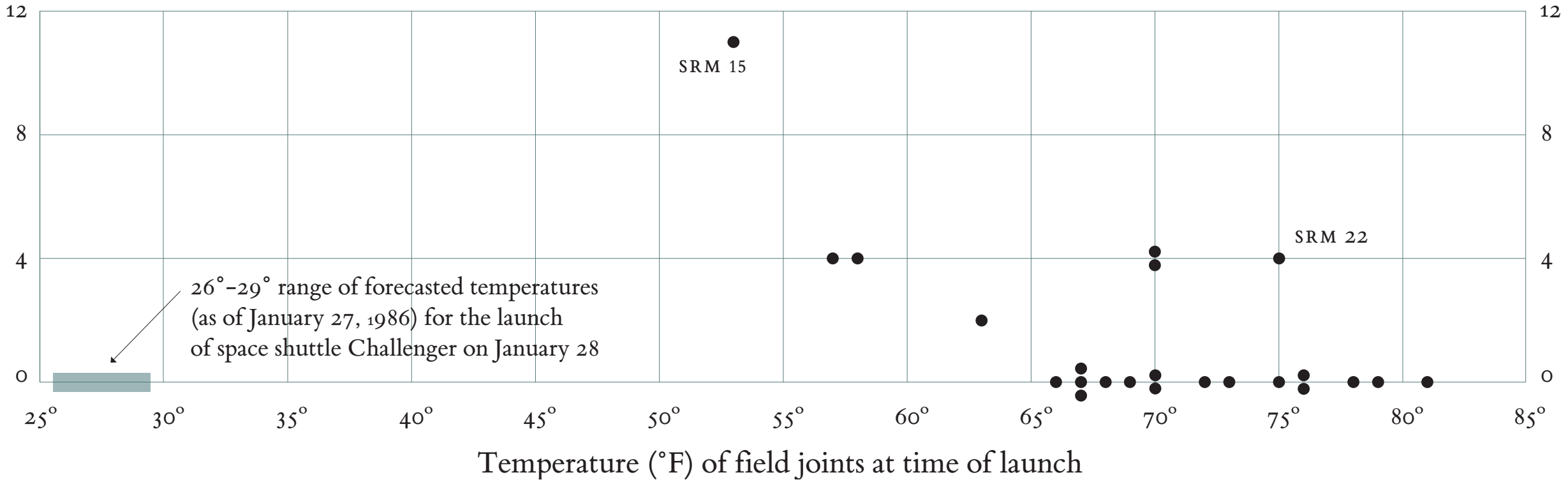
The flight crew of Challenger 51-L. Front row, left to right: Michael J. Smith, pilot; Francis R. (Dick) Scobee, commander; Ronald E. McNair. Back row: Ellison S. Onizuka, S. Christa McAuliffe, Gregory B. Jarvis, Judith A. Resnik.

On January 28, 1986 Challenger broke apart 73 seconds into its flight, leading to the deaths of its seven crew members.



What caused the failure?

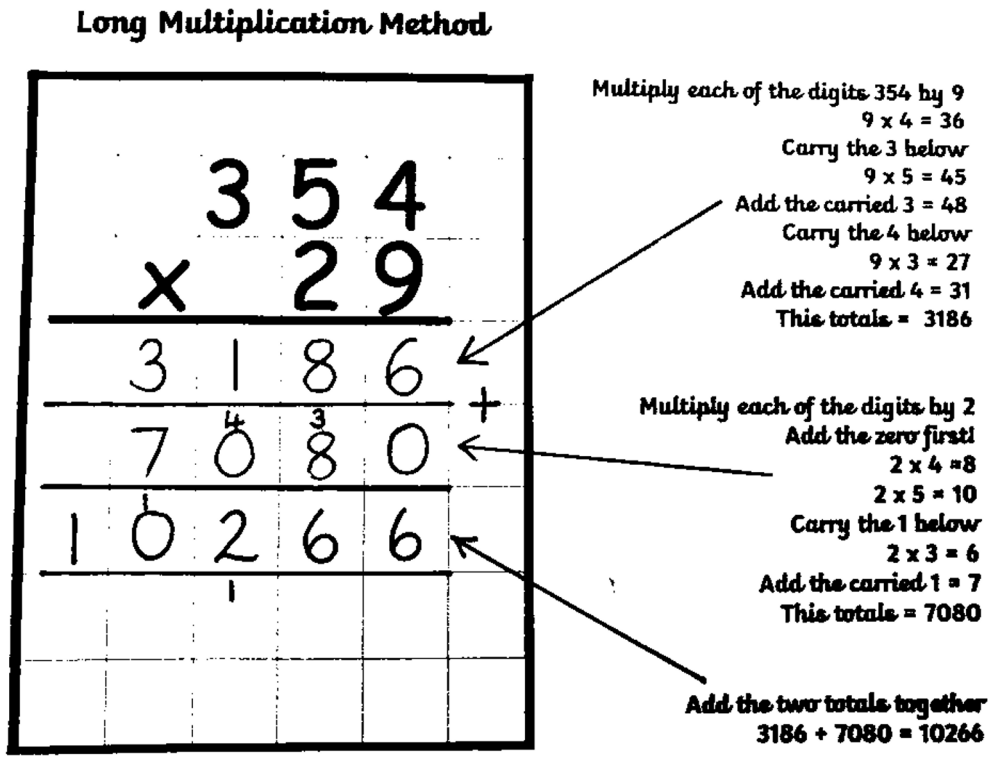
Support Reasoning

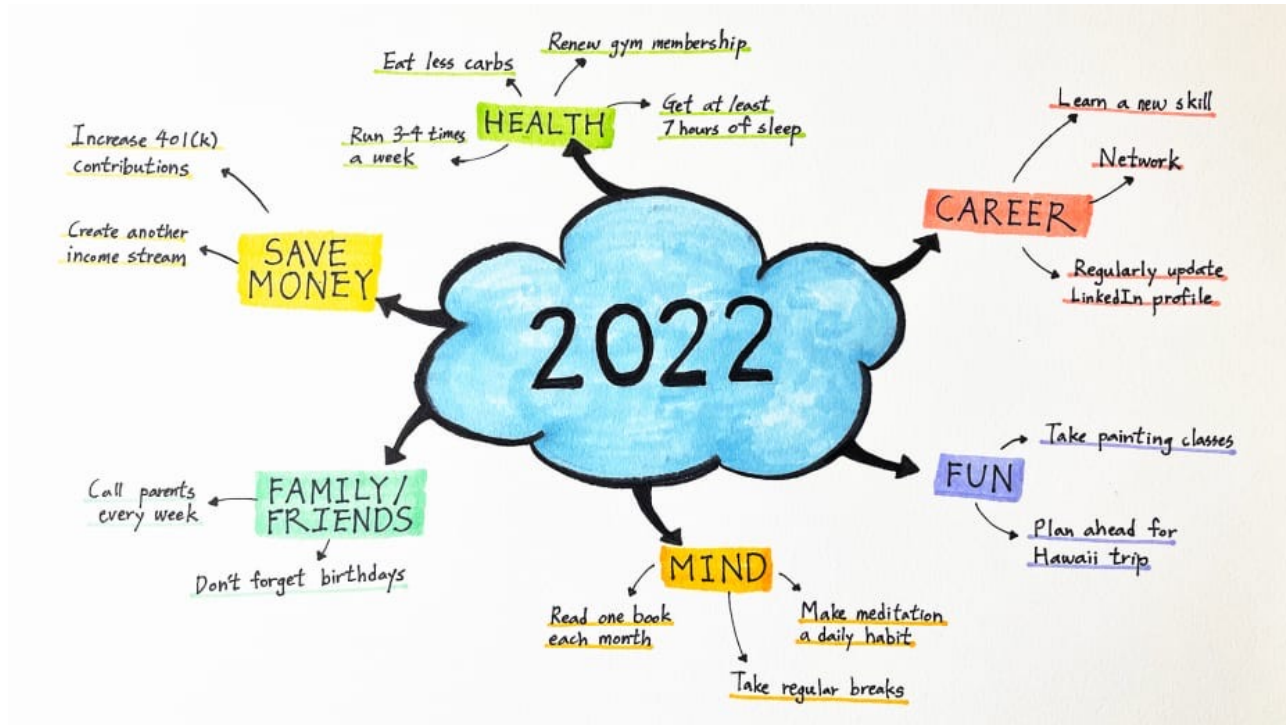


What caused the failure? Chilly temperatures on launch day

What is a damage index?
Which temperatures? O-ring?
Outside?

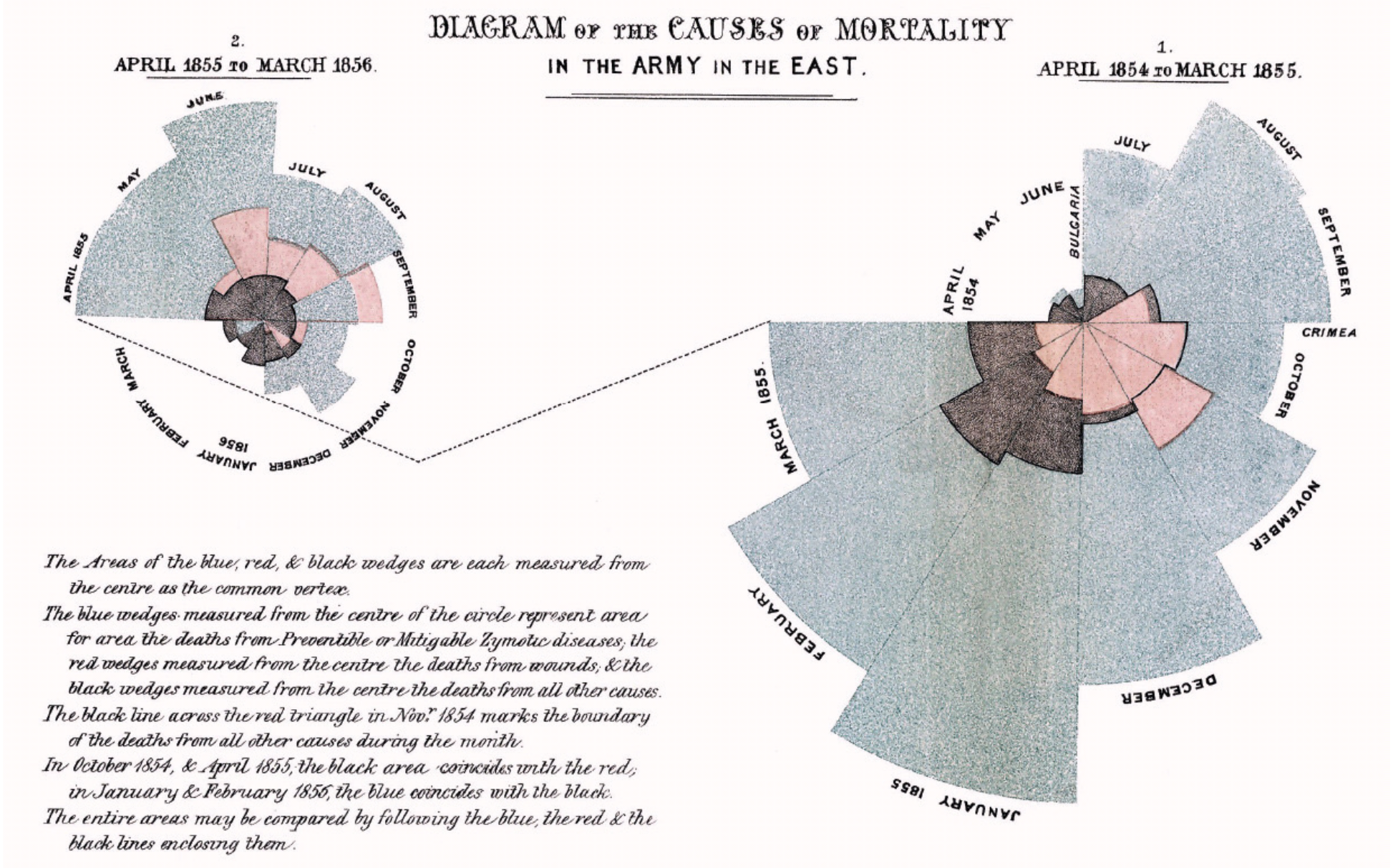
Support Reasoning





Support Reasoning

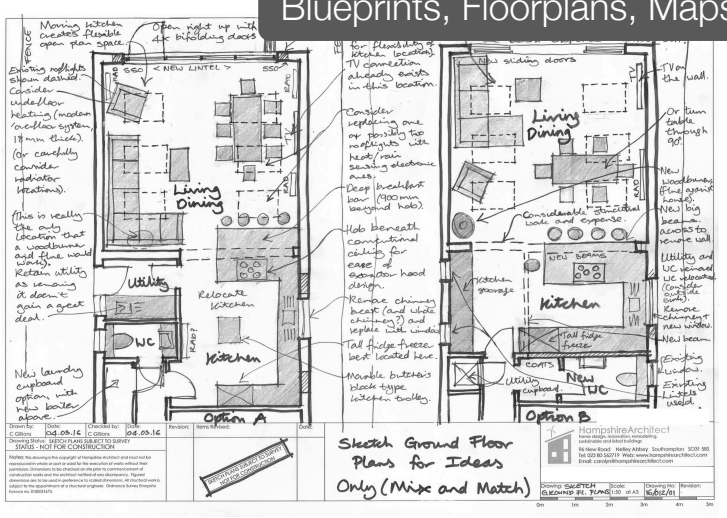
Convey Information



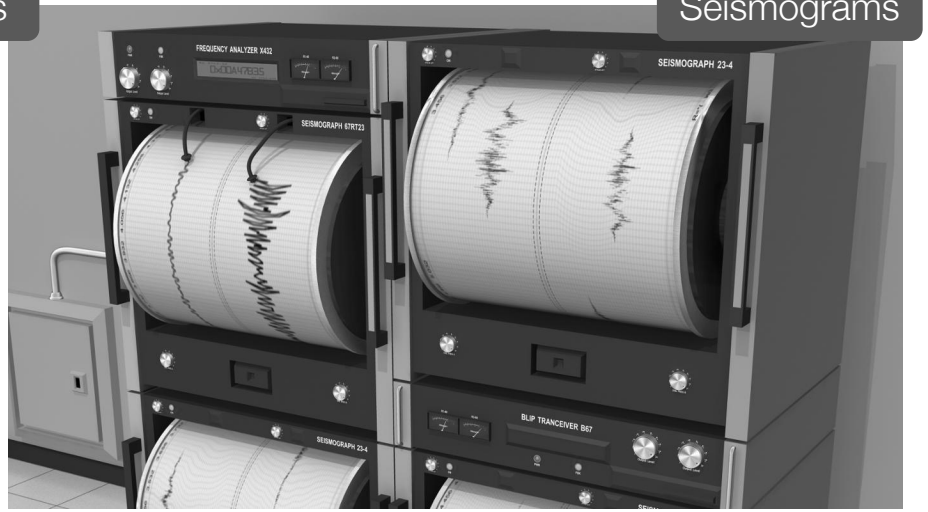
Florence Nightingale's 1856 Crimean War Coxcombs: "to affect thro' the Eyes what we fail to convey to the public through their word-proof ears"

Convey Information

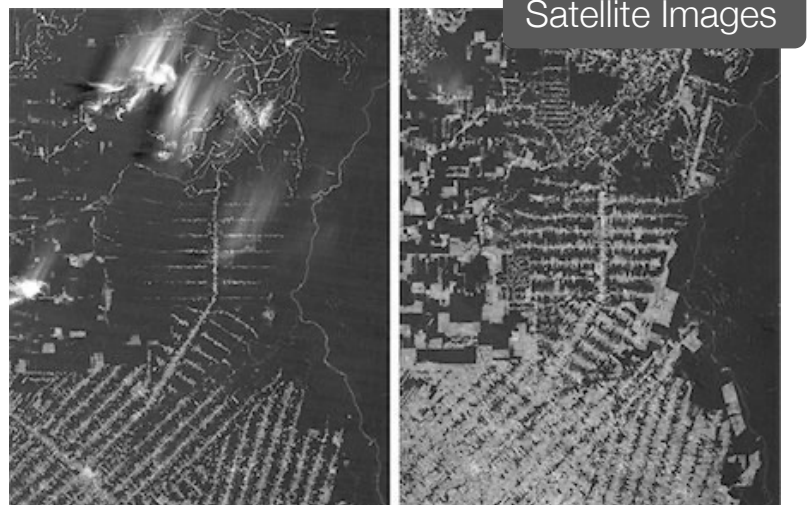
Blueprints, Floorplans, Maps



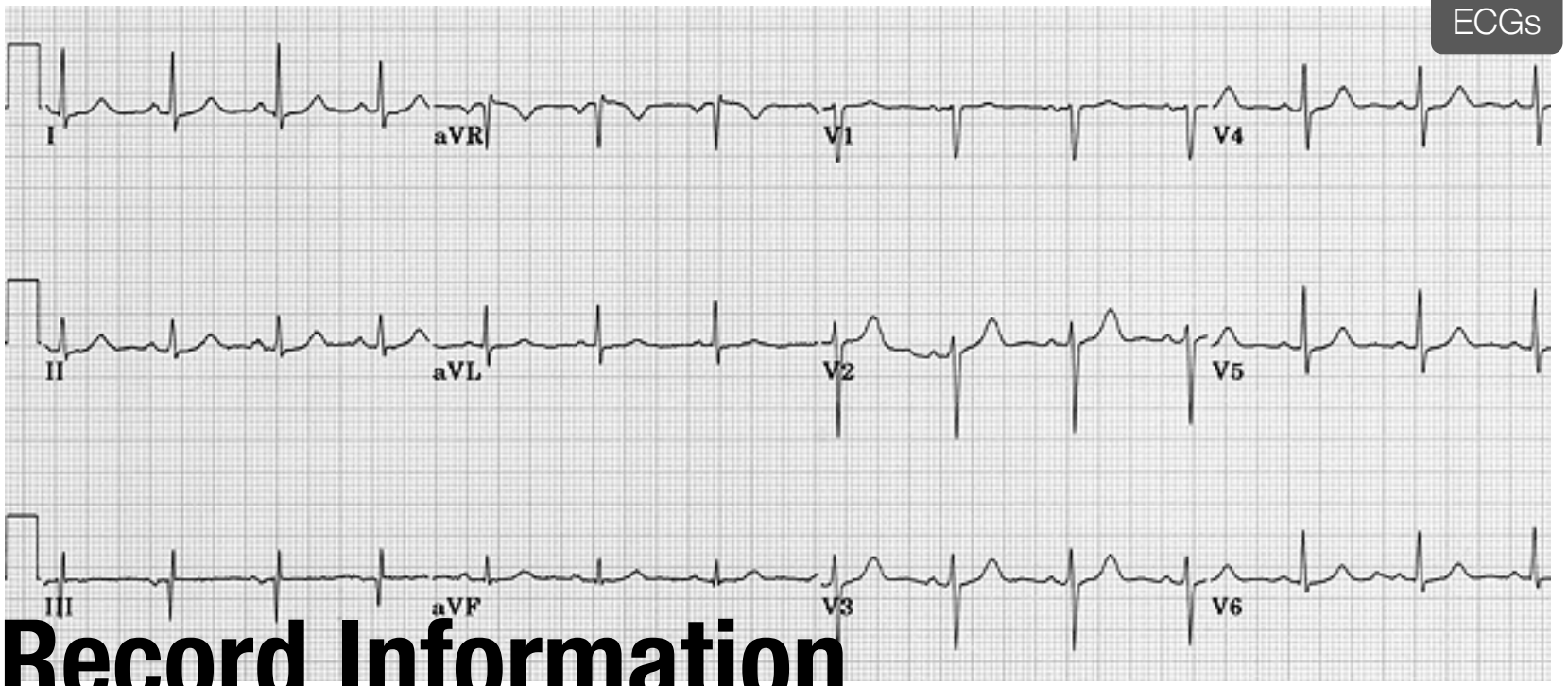
Seismograms



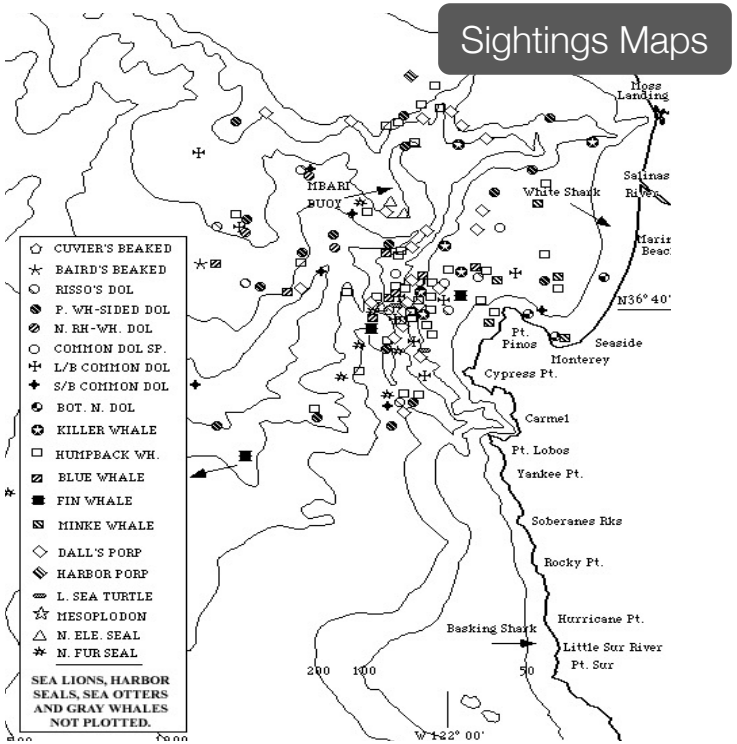
Satellite Images



ECGs



Sightings Maps



Record Information

How to make good visualizations?

vis·u·al·ize

/ˈviZH(o͞o)əˌlīz/

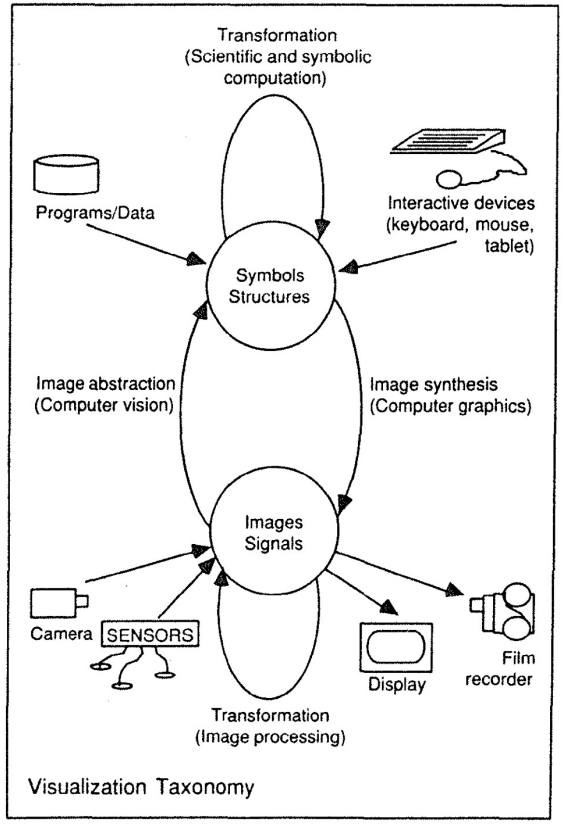
verb

1. form a mental image of; imagine.
   "it is not easy to visualize the future"

   Similar: envisage | envision | conjure up

2. make (something) visible to the eye.
   "the cells were better visualized by staining"

Definitions from Oxford Languages



Symbolic

↓

Transformation

↓

Geometric

McCormick et al, Visualization in Scientific Computing, Computer Graphics 21, 6 (November 1987)

Gives an operational definition of visualization:
It is the visual encoding of data variables

What is a visualization?

Design Criteria

Expressiveness

Tell the truth, the whole truth and nothing but the truth!

A set of facts is expressible in a visual language if the sentences (i.e. the visualizations) in the language **express** all **the facts** in the set of data, and only the facts in the data.

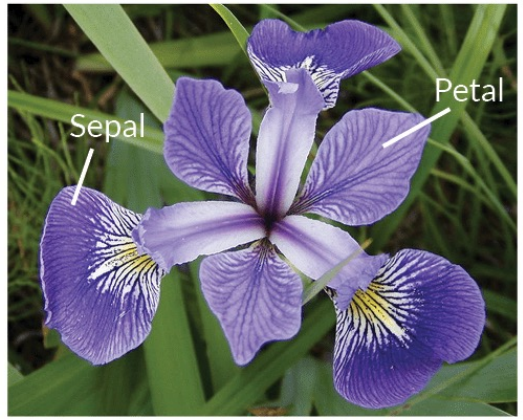
Effectiveness

Use visual encodings that people can decode better (faster and more accurate).

A visualization is more effective than another visualization if the information conveyed by one visualization is more readily perceived than the information in the other visualization.

McKinlay, 1986

Expressiveness

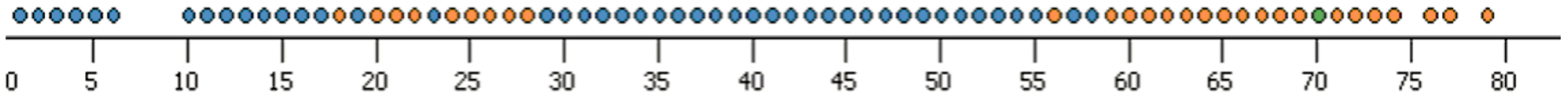


Iris Versicolor

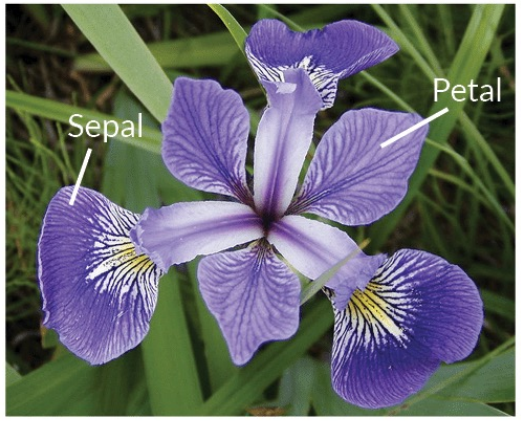
Iris Setosa

Iris Verginica

Expressive?



Sepal and petal area

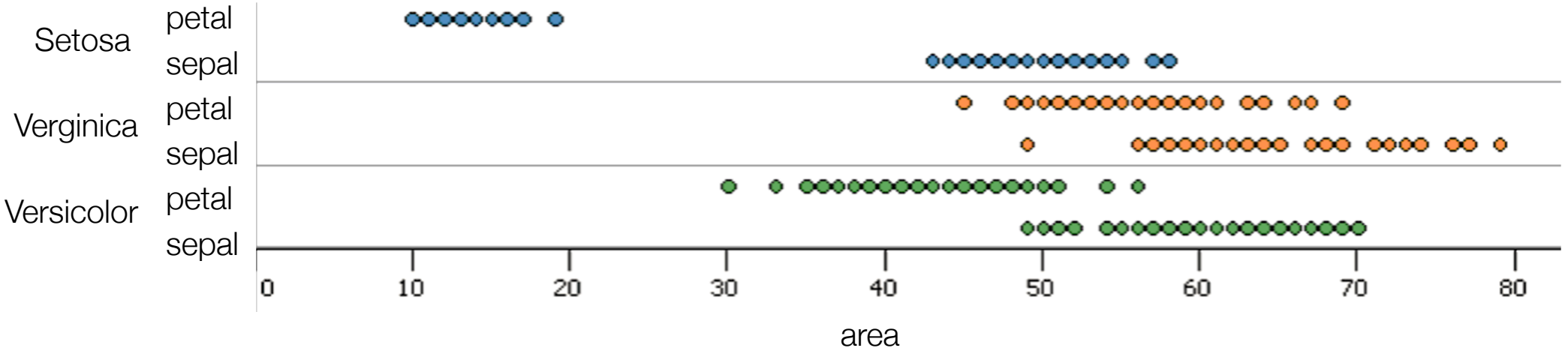


Iris Versicolor

Iris Setosa

Iris Verginica

Expressive?



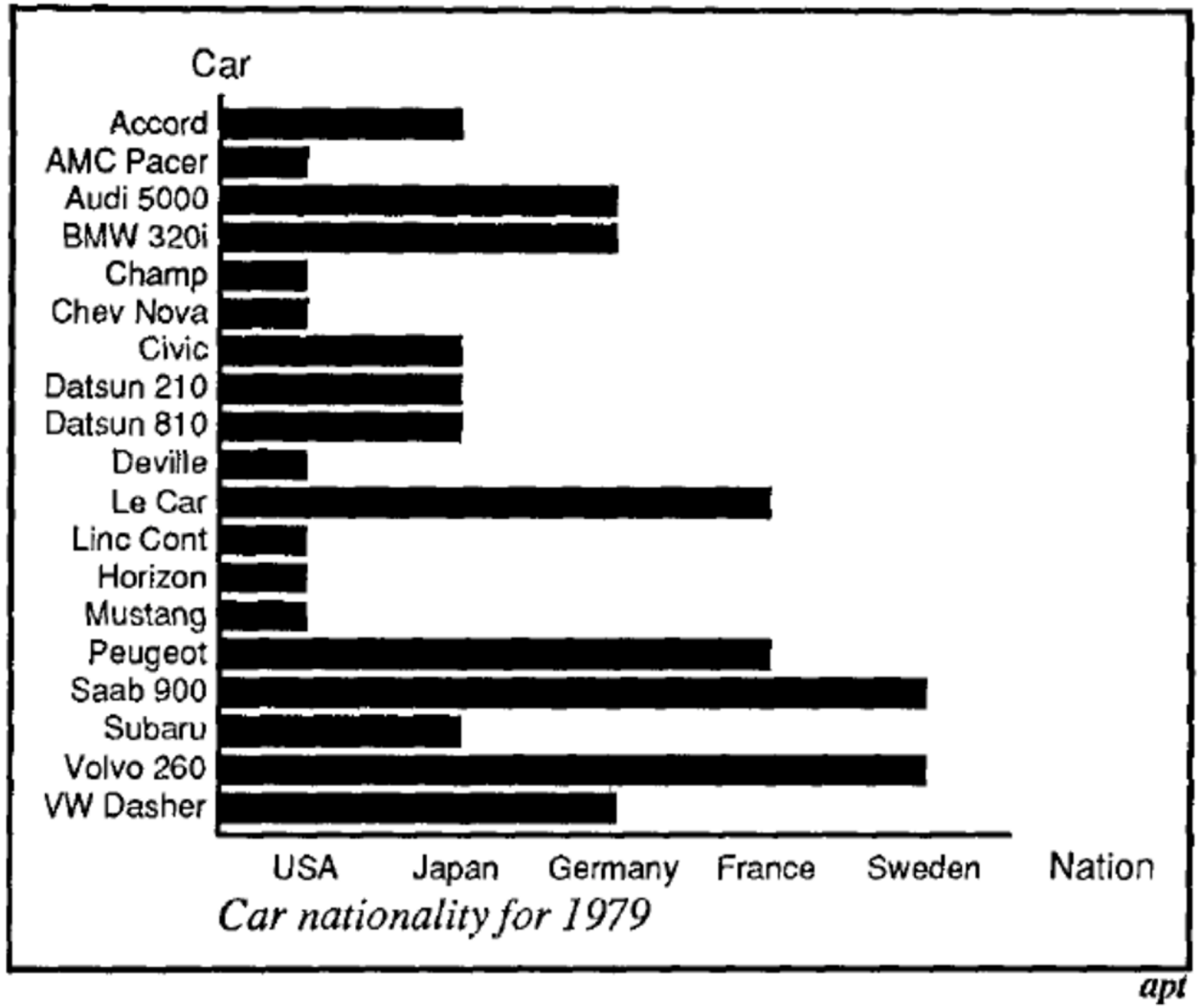


Fig. 11. Incorrect use of a bar chart for the *Nation* relation. The lengths of the bars suggest an ordering on the vertical axis, as if the USA cars were longer or better than the other cars, which is not true for the *Nation* relation.

Expressive?

Effectiveness
“Use visual encodings that people can decode better”

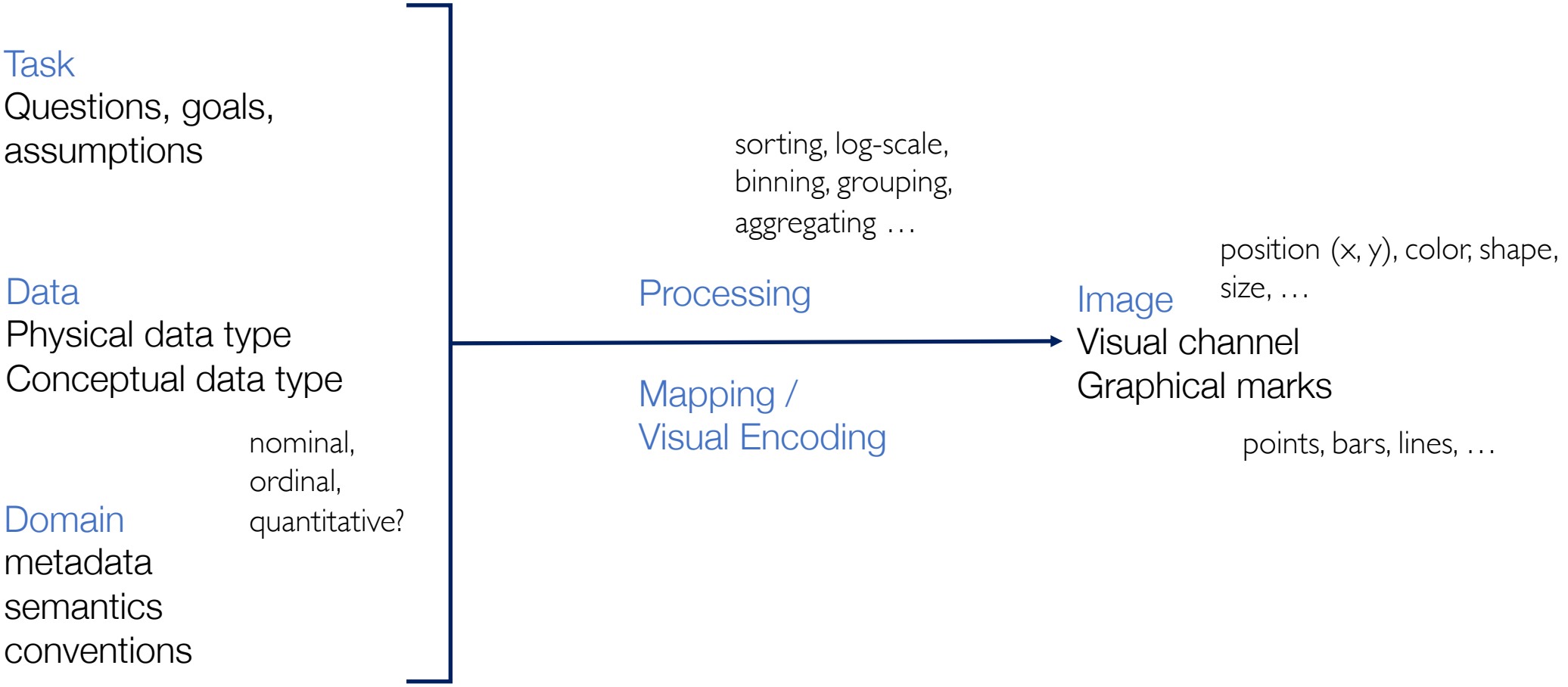
1. Learn about Data Models

2. Learn about Visual Encoding

3. Learn the Rules

4. Examples

5. Summarize with practical tips



Outline