Visualisation

Hugh Shanahan
hugh.Shanahan@rhul.ac.uk
@hughshanahan

Why visualise?

We have Excel!!!

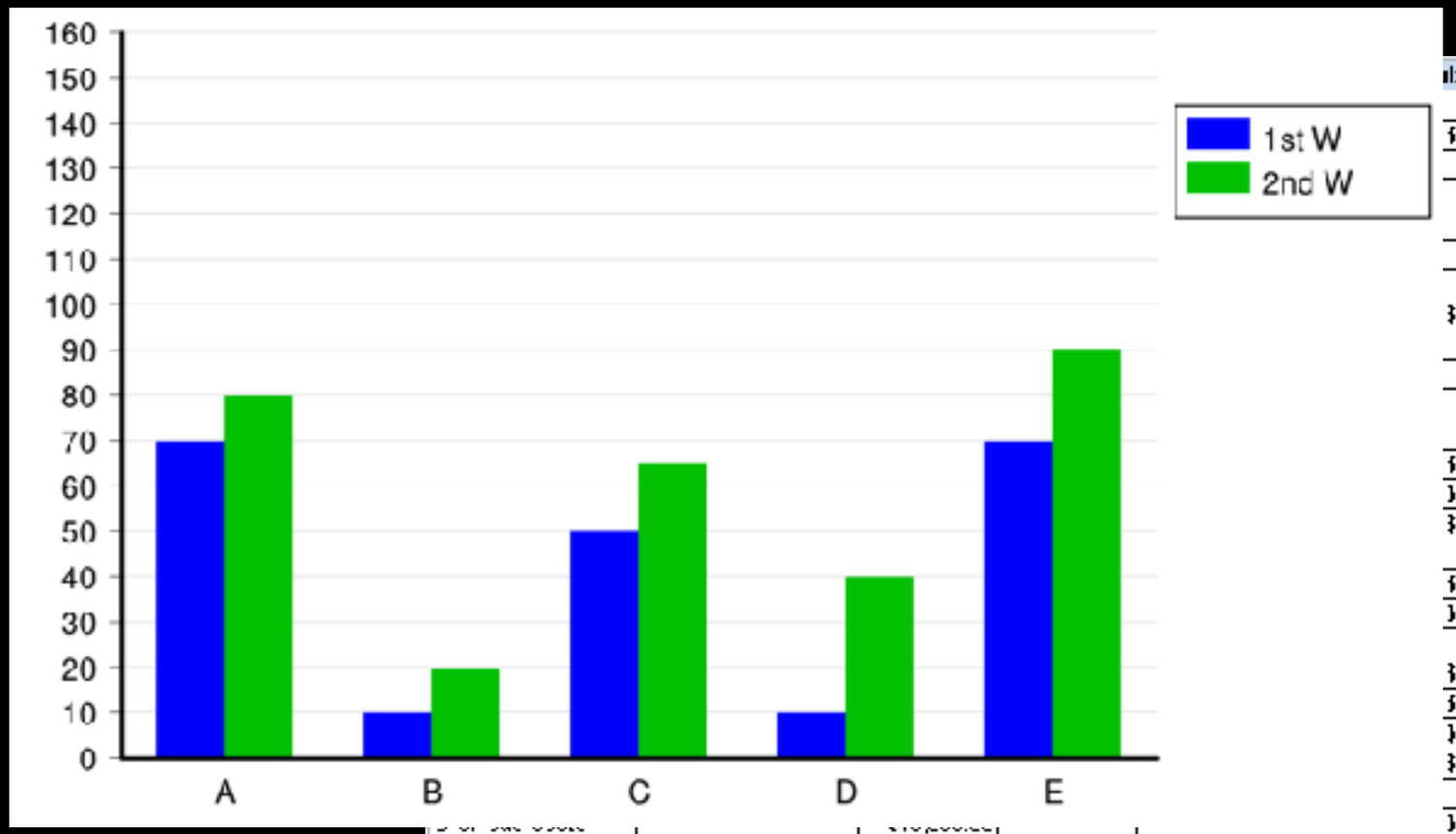
Why visualise?

We have Excel!!!

Work Item	Vendor	Labor	Equipment	Materials	Subcontr.	Subtotal	Markup %	Markup	Total
Permits/Fees	City of Los Angeles				$1,500.00	$1,500.00		$0.00	$1,500.00
Excavation		$6,000.00	$8,000.00	$400.00		$14,400.00	15.00%	$2,175.00	$16,575.00
Utilities		$3,200.00	$2,300.00	$2,700.00	$1,000.00	$9,750.00	10.00%	$1,462.50	$11,212.50
Water Well						$0.00		$0.00	$0.00
Septic Tank						$0.00		$0.00	$0.00
Foundation	Corral's Concrete				$3,500.00	$3,500.00	5.00%	$175.00	$3,675.00
Concrete Flatwork	Corral's Concrete				$1,900.00	$1,900.00	5.00%	$95.00	$1,995.00
Framing		$3,500.00	$1,500.00	$8,000.00		$14,000.00	15.00%	$2,100.00	$16,100.00
Roofing	Patrick's Roofing				$5,000.00	$5,000.00	5.00%	$175.00	$5,175.00
Windows/Ext Doors	Wally's Windows				$3,000.00	$3,000.00	5.00%	$150.00	$3,150.00
Garage Door	Gary's Garage Doors				$2,250.00	$2,250.00	5.00%	$112.50	$2,362.50
Siding						$0.00		$0.00	$0.00
Electrical	Frank's Electric				$15,000.00	$15,000.00	5.00%	$825.00	$15,825.00
Plumbing	Mac's Mechanical				$15,000.00	$15,000.00	5.00%	$825.00	$15,825.00
HVAC	Mac's Mechanical				$23,000.00	$23,000.00	5.00%	$1,150.00	$24,150.00
Insulation		$3,500.00		$1,000.00		$4,500.00		$0.00	$4,500.00
Masonry	Marcus Masonry				$14,000.00	$14,000.00	5.00%	$700.00	$14,700.00
Drywall	Doug's Drywall				$12,000.00	$12,000.00	5.00%	$600.00	$12,600.00
Interior Trim	Doug's Drywall				$3,000.00	$3,000.00	5.00%	$150.00	$3,150.00
Painting	Paul's Painting				$13,500.00	$13,500.00	5.00%	$675.00	$14,175.00
Floor Coverings	Carl's Carpets				$16,500.00	$16,500.00	5.00%	$825.00	$17,325.00
Cabinets	Ken's Cabinets				$22,500.00	$22,500.00	5.00%	$1,125.00	$23,625.00
Appliances	Abby's Appliances	$2,500.00		$11,500.00		$14,000.00	15.00%	$2,100.00	$16,100.00
Landscaping	Sonny's Sodding				$2,750.00	$2,750.00	5.00%	$137.50	$2,887.50
Overhead Costs		$10,000.00				$10,000.00	20.00%	$2,000.00	$12,000.00
Other						$0.00		$0.00	$0.00
						$0.00		$0.00	$0.00
						$0.00		$0.00	$0.00
						$0.00		$0.00	$0.00
						$0.00		$0.00	$0.00
						$0.00		$0.00	$0.00
TOTALS					$236,650.00		7.71%	$18,257.50	$254,907.50

Why visualise?

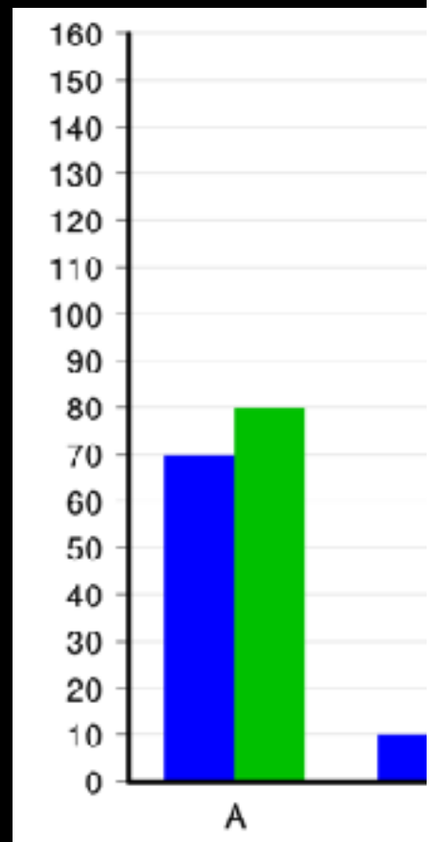
We have Excel!!!

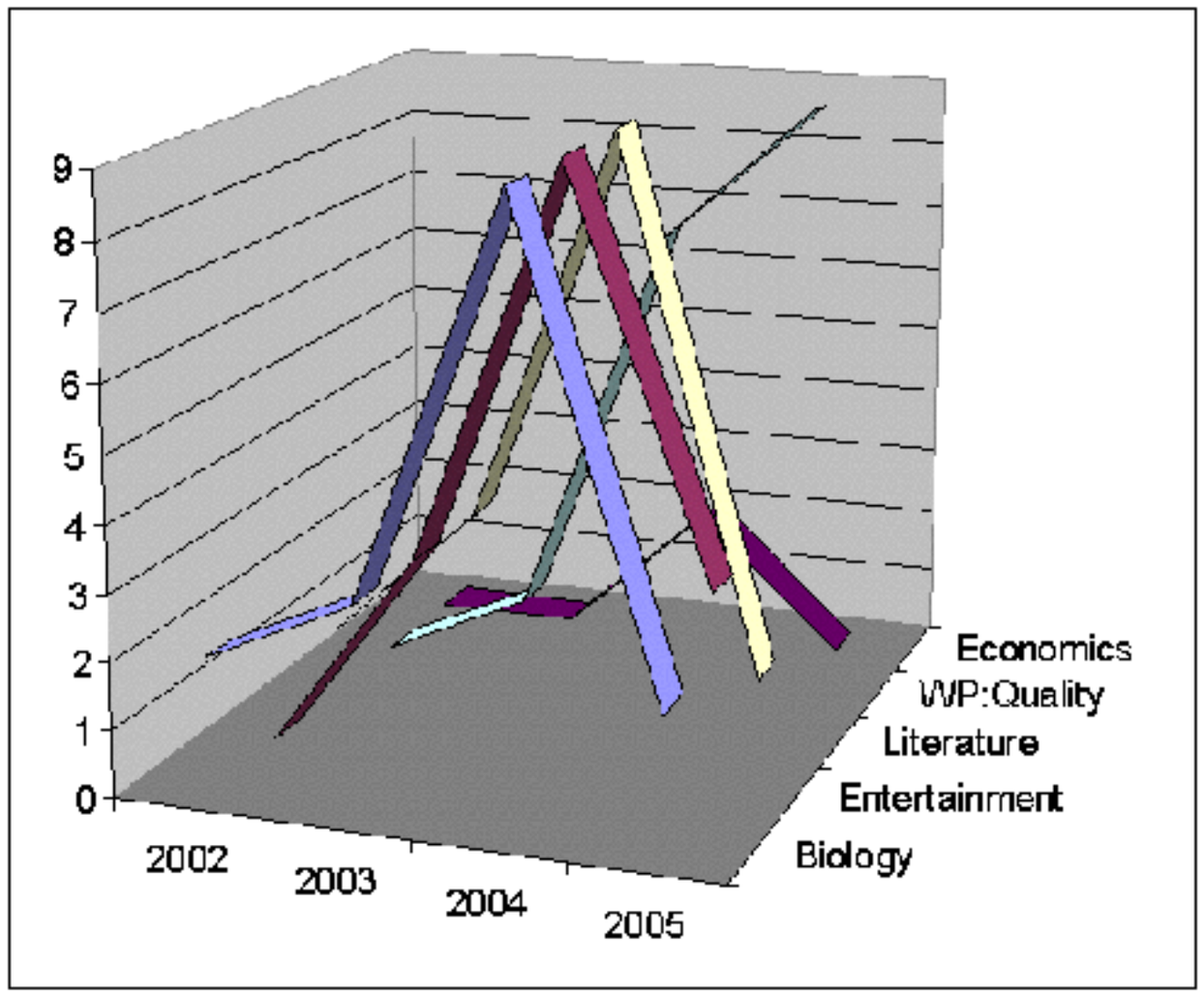


				ubtotal	Markup %	Markup	Total
				$[illegible]		$0.00	$[illegible]
				$14 [illegible]	15 [illegible]%	$2 [illegible]	$16 [illegible]
				$9 750.00	10 00%	$[illegible]	$[illegible]
				$0.00		$0.00	$0.00
				$[illegible]		$[illegible]	$[illegible]
				$3 500.00	5 00%	$175.00	$3 675.00
				$[illegible]	5 00%	$[illegible]	$[illegible]
				$14 000.00	15 00%	$2 100.00	$16 100.00
				$[illegible]	5 [illegible]%	$[illegible]	$[illegible]
				$[illegible]	5 00%	$[illegible]	$[illegible]
				$2 250.00	5 00%	$[illegible]	$2 362.50
				$0.00		$0.00	$0.00
				$[illegible]	5 [illegible]%	$[illegible]	$[illegible]
				$16 500.00	5 00%	$825.00	$17 325.00
				$23 000.00	5 00%	$[illegible]	$24 150.00
				$4 [illegible]		$[illegible]	$4 [illegible]
				$14 [illegible]	5 [illegible]%	$[illegible]	$[illegible]
				$12 500.00	5 00%	$625.00	$13 125.00
				$3 000.00	5 00%	$450.00	$3 450.00
				$13 500.00	5 00%	$675.00	$14 175.00
				$16 [illegible]	5 00%	$825.00	$17 325.00
				$22 500.00	5 00%	$[illegible]	$23 625.00
				$14 000.00	15 00%	$2 100.00	$16 100.00
				$2 [illegible]	5 [illegible]%	$[illegible]	$2 [illegible]
				$12 000.00	20 00%	$2 000.00	$12 000.00
Other				$0.00		$0.00	$0.00
				$[illegible]		$[illegible]	$[illegible]
				$[illegible]		$[illegible]	$[illegible]
				$0.00		$0.00	$0.00
				$0.00		$0.00	$0.00
				$[illegible]		$[illegible]	$[illegible]
TOTALS				$236,650.00	7.71%	$18,257.50	$254,907.50

Why visualise?

We have E



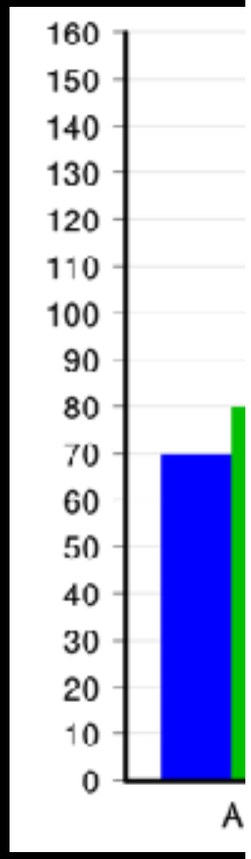


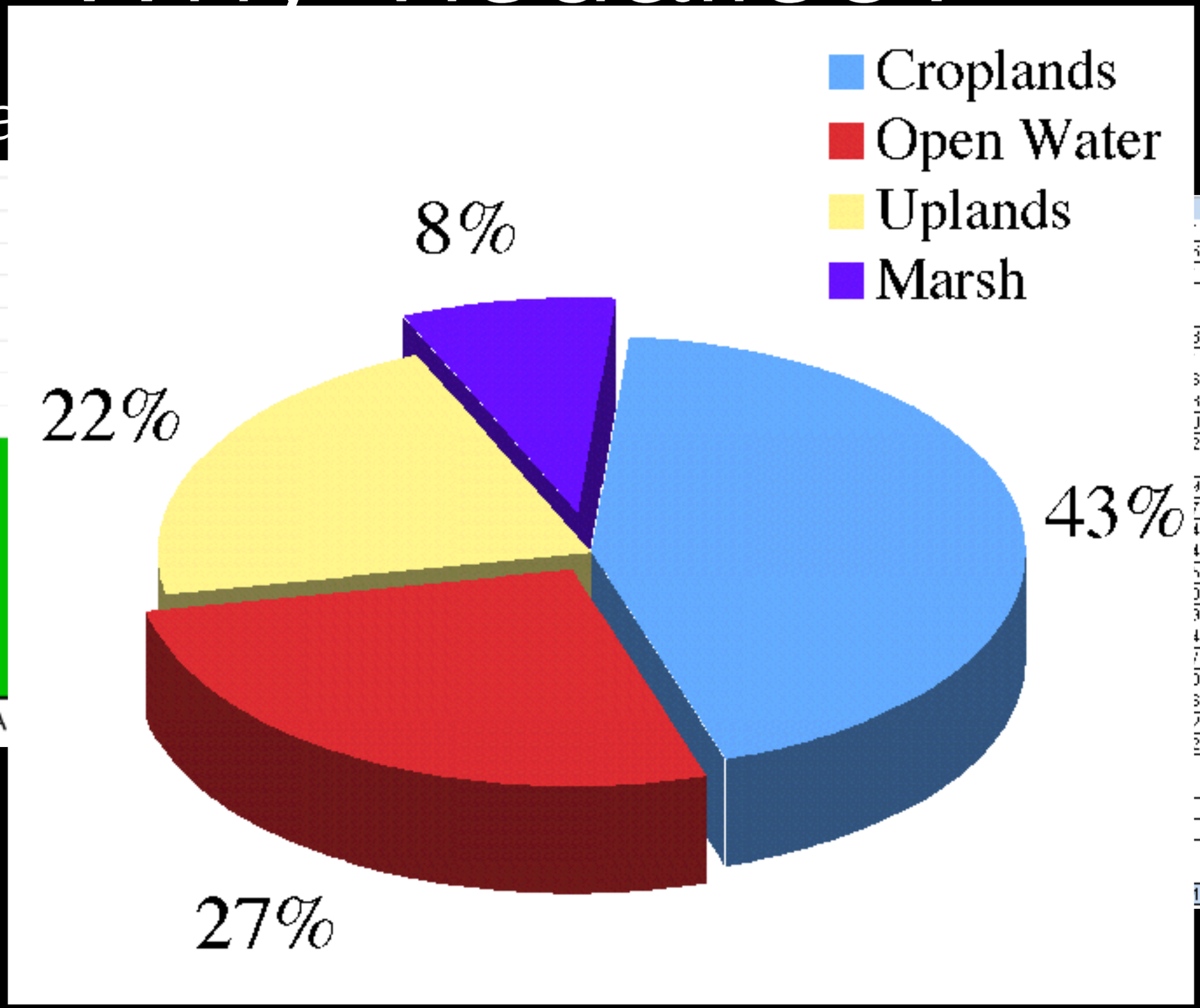
				$0.00	$0.00	$0.00
				$7.00	$7.00	$7.00
TOTALS				$236,600.00	7.71%	$18,257.50	$254,907.50

Assorted images from Wikipedia

Why visualise?

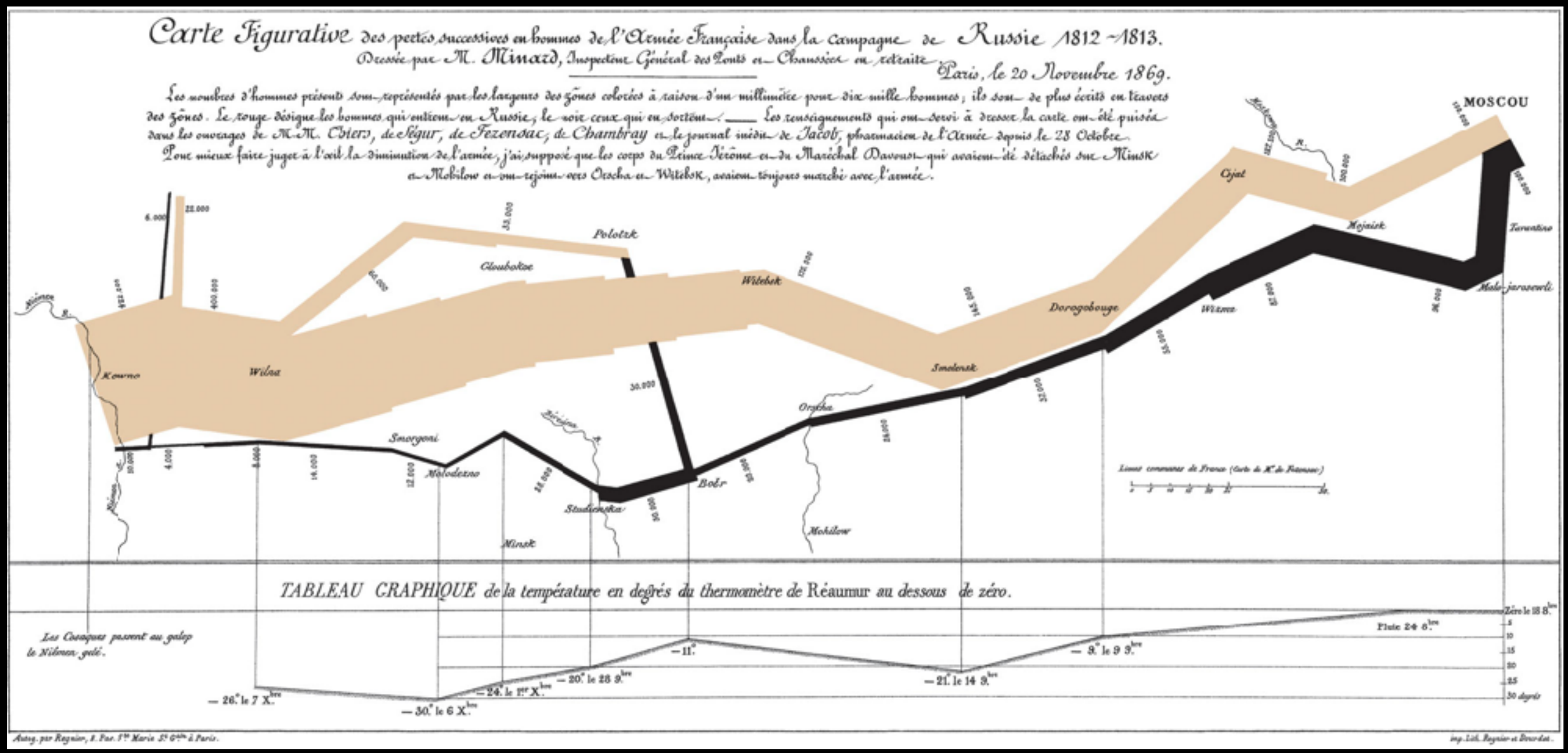
We ha





Napoleonic 1812 campaign to Russia

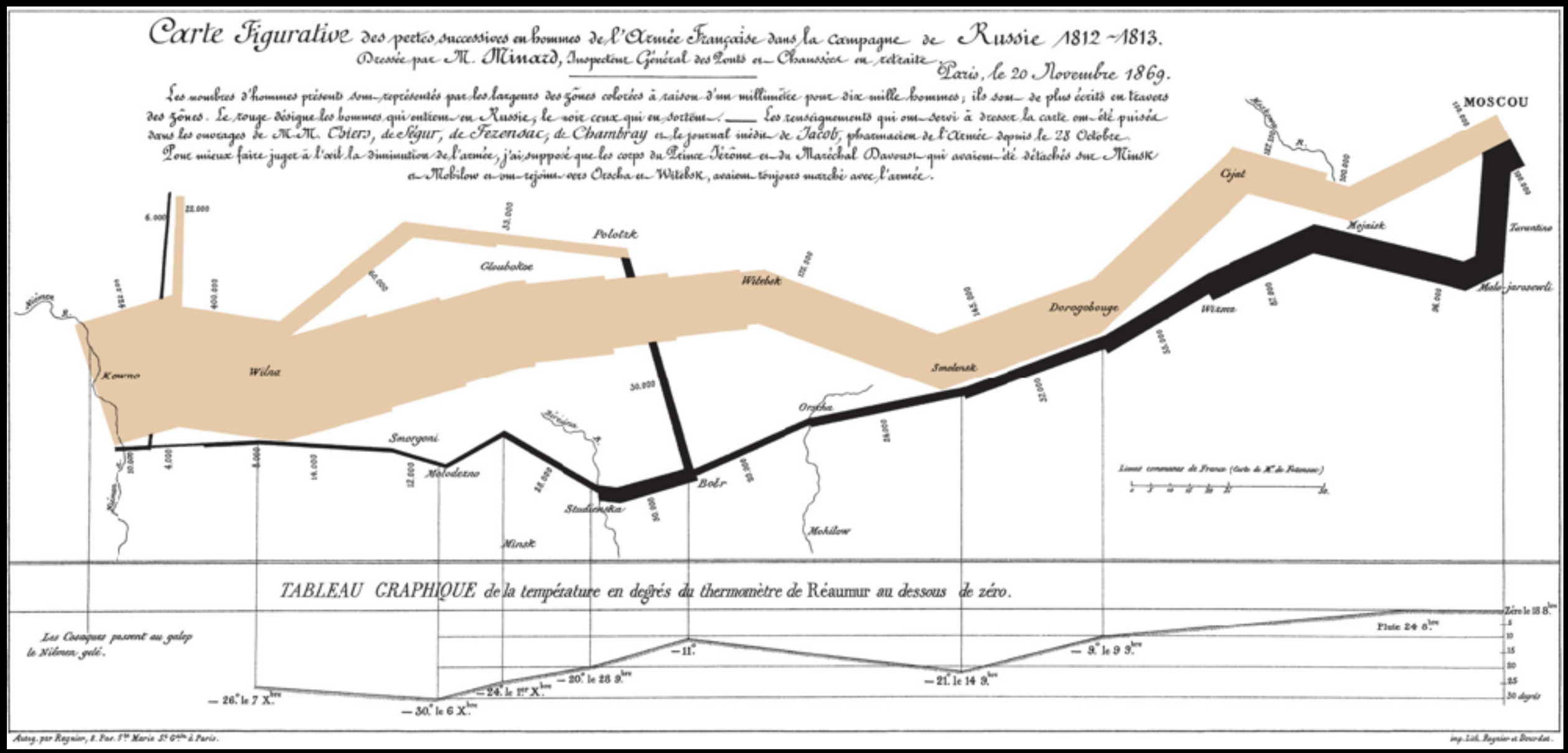
Napoleonic 1812 campaign to Russia



https://en.wikipedia.org/wiki/File:Minard.png

Napoleonic 1812 campaign to Russia

Charles Minard 1869



https://en.wikipedia.org/wiki/File:Minard.png

What's Visualisation and why do it?

Visualisation (Vis) allows people to analyse data when they don't know exactly what question to ask in advance

https://www.youtube.com/watch?v=b5C5dV9XVKo

WARNING - CONTAINS DRAWINGS OF TESTICLES, RECTAL BLEEDING, VOMITING AND DONALD TRUMP

Visualisation

Not just a bunch of standard boring plots

A tool to help

You understand the data

To explain the data

And now the hard bit…

We need a tool to do visualisations

And now the hard bit…

We need a tool to do visualisations

Easy
Inflexible

Hard
Flexible

And now the hard bit…

We need a tool to do visualisations

Easy
Inflexible

Hard
Flexible

Tableau

Excel

Plotly

And now the hard bit…

We need a tool to do visualisations

Easy
Inflexible

Hard
Flexible

Tableau

Excel

D3

Plotly

And now the hard bit…

We need a tool to do visualisations

Easy
Inflexible

Hard
Flexible

Tableau

Excel

R (ggplot)
Python
(matplotlib,seaborn)

D3

Plotly

What's Visualisation and why do it?

Visualisation (Vis) allows people to analyse data when they don't know exactly what question to ask in advance

Based on Chapter 1 of "Visualisation, Analysis and Design"

Back one step…

What do we mean by data?

Back one step…

What do we mean by data?

More complex than one might think!

Back one step…

What do we mean by data?

More complex than one might think!

Big Data, Little Data, No Data, Christine Borgman

Back one step…

What do we mean by data?

More complex than one might think!

Big Data, Little Data, No Data, Christine Borgman

“Data” includes, at a minimum, digital observation, scientific monitoring, data from sensors, metadata, model output and scenarios, qualitative or observed behavioural data, visualizations, and statistical data collected for administrative or commercial purposes. Data are generally viewed as input to the research process.

Fox and Harris Data Science Journal Vol 12, 10 Feb 2013

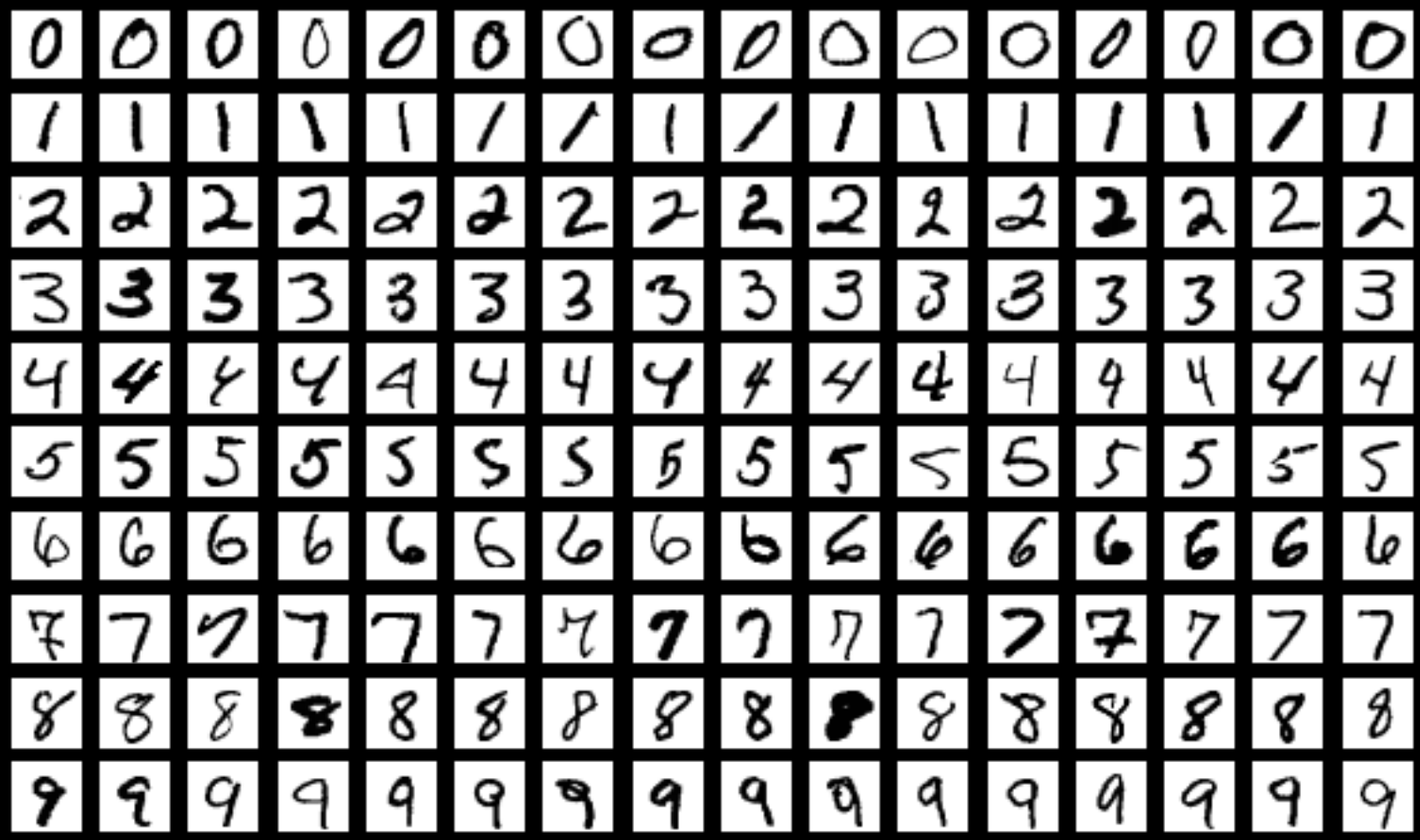
More colloquially

Data (for visualisation) is that which can be stored electronically….

Sometimes don’t need Vis

Many cases where a statistical analysis or using Machine Learning is sufficient

Example MNIST Data Set



By Josef Steppan - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=64810040

But often you do need Vis

But often you do need Vis

analysis problems are ill specified

But often you do need Vis

analysis problems are ill specified

people don’t know how to approach the problem

But often you do need Vis

analysis problems are ill specified

people don’t know how to approach the problem

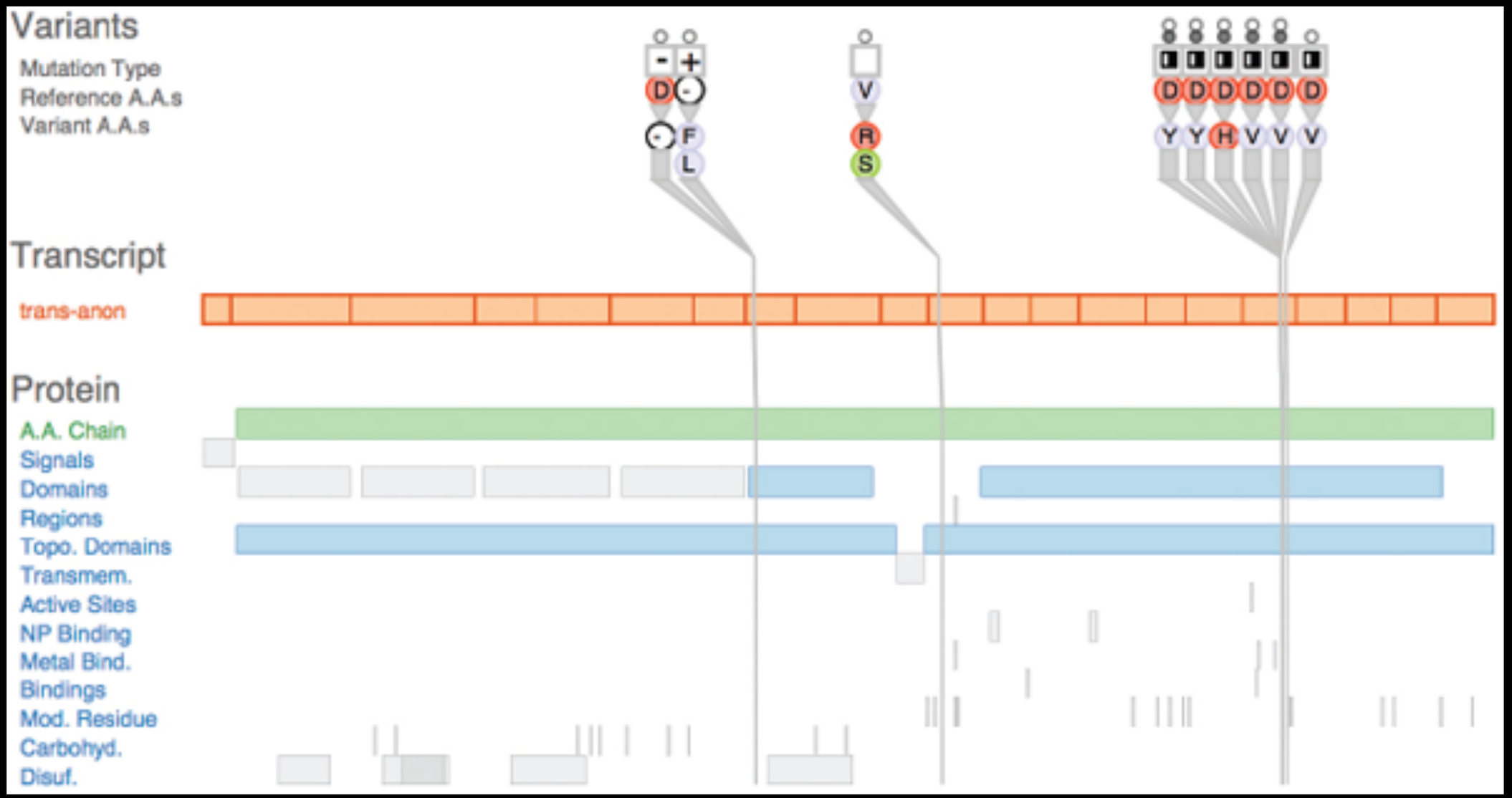
Augment human capabilities rather than replace humans

Vis can be..

For long term use - visualisations that are updated and used again and again

For short term/one off use - understanding something about the data; telling a story

Long term use



Genomic sequences - showing variation

Joel A. Ferstay, Cydney B. Nielsen, and Tamara Munzner. "Variant View: Visualizing Sequence Variants in Their Gene Context." *IEEE Transactions on Visualization Computer Graphics (Proc. InfoVis 13)* 19:12 (2013), 2546–2555. *(page 4)*

Why Computational tools?

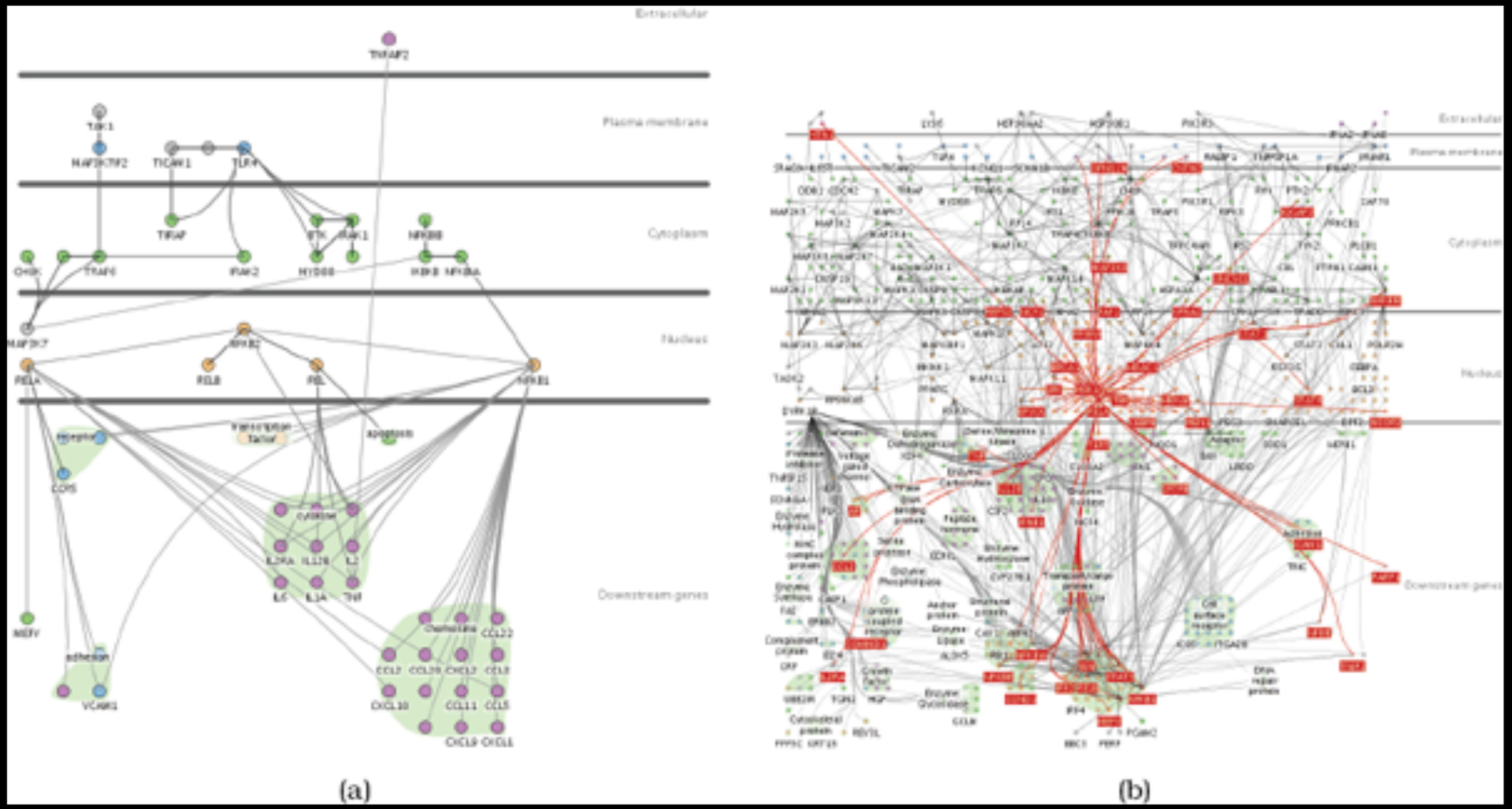
Mona Chalabi and Charles Minard did it by hand!

Need to visualise

Large/Complex data sets

Variable data sets

Quickly and in an exploratory fashion



Could be done by hand

Too hard to do by hand!

Aaron Barsky, Jennifer L. Gardy, Robert E. Hancock, and Tamara Munzner. “Cerebral: A Cytoscape Plugin for Layout of and Interaction with Biological Networks Using Subcellular Localization Annotation.” *Bioinformatics* 23:8 (2007), 1040–1042. *(pages 5, 6, 295)*

Why Vision?

Able to absorb large amounts of information in parallel.

Why Vision?

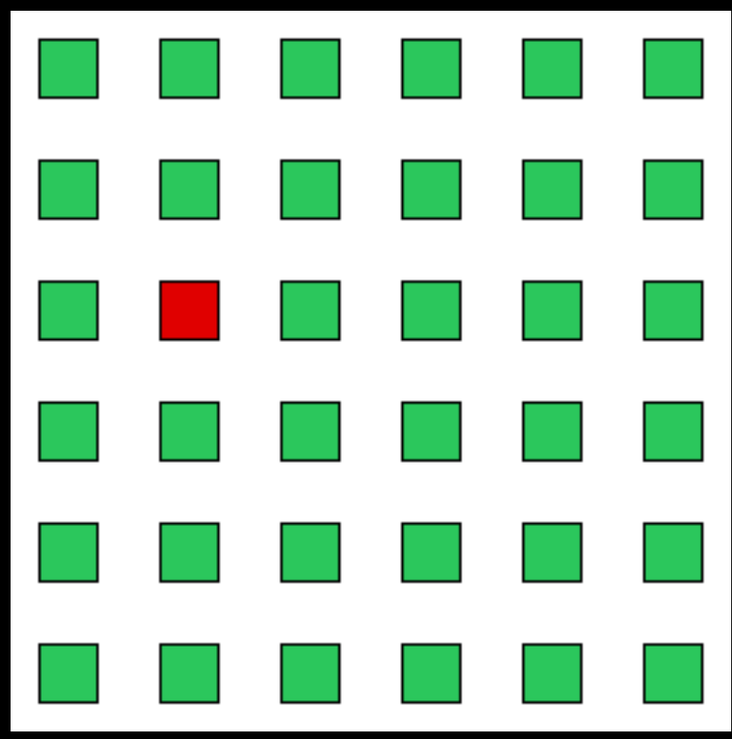
Able to absorb large amounts of information in parallel.

Example - Visual Pop-out

Why Vision?

Able to absorb large amounts of information in parallel.

Example - Visual Pop-out



By Head 21:20, 24 August 2007 (UTC) (self-drawn) [Public domain], via Wikimedia Commons

Other senses?

Sound - sequence-based, not parallel

Listen to individual voices in a choir (for example)

Why do we need to look at data in detail?

Why do we need to look at data in detail?

Even “simple” data can mislead us

Why do we need to look at data in detail?

Even “simple” data can mislead us

Need to check if summary statistics are right.

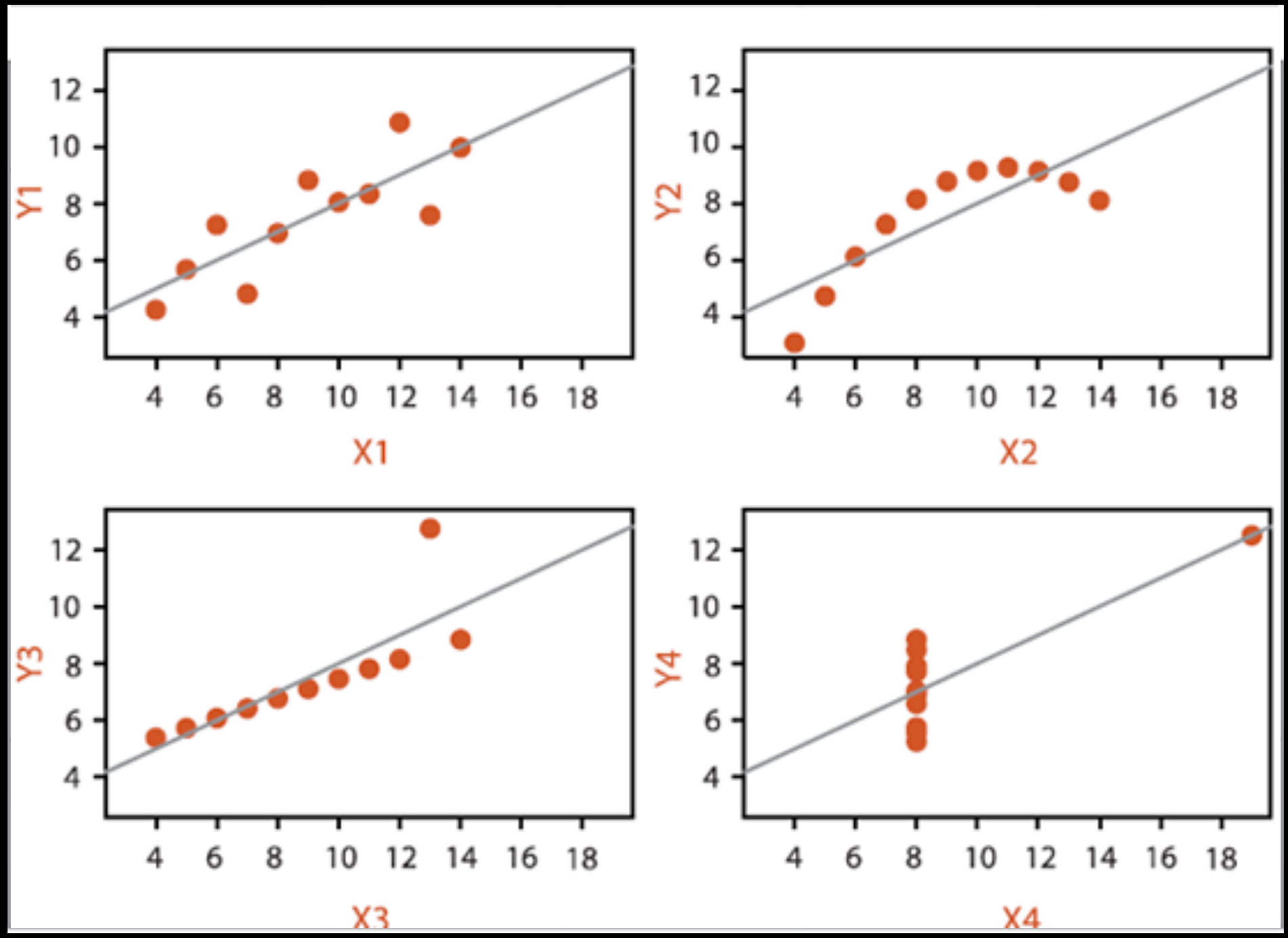
Anscombe's Quartet

Anscombe's Quartet: Raw Data

	1		2		3		4	
	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
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	

F.J. Anscombe. "Graphs in Statistical Analysis." *American Statistician* 27 (1973), 17–21. *(pages 7, 8, 19)*

Anscombe's Quartet



F.J. Anscombe. "Graphs in Statistical Analysis." *American Statistician* 27 (1973), 17–21. *(pages 7, 8, 19)*

Interaction

Interaction

Same data set - many different ways to visualise

Interaction

Same data set - many different ways to visualise

Tools to analyse particular type of data

Interaction

Same data set - many different ways to visualise

Tools to analyse particular type of data

Variant View - discussed above

Interaction

Same data set - many different ways to visualise

Tools to analyse particular type of data

Variant View - discussed above

But can also interact by writing new code

Vis Idiom

A Vis idiom is a distinct way to creating and manipulating visual representations.

Vis Idiom

A Vis idiom is a distinct way to creating and manipulating visual representations.

Examples of idioms

Scatter plots

Bar plots

Pie Charts (please don’t)

Vis Idiom

A Vis idiom is a distinct way to creating and manipulating visual representations.

Examples of idioms

Scatter plots

Bar plots

Pie Charts (please don’t)

Huge possible number of Idioms

See for example Information is Beautiful

Tasks

No one visualisation does everything

Task - what is the purpose of the visualisation?

Journalism or telling a story?

Exploring data?

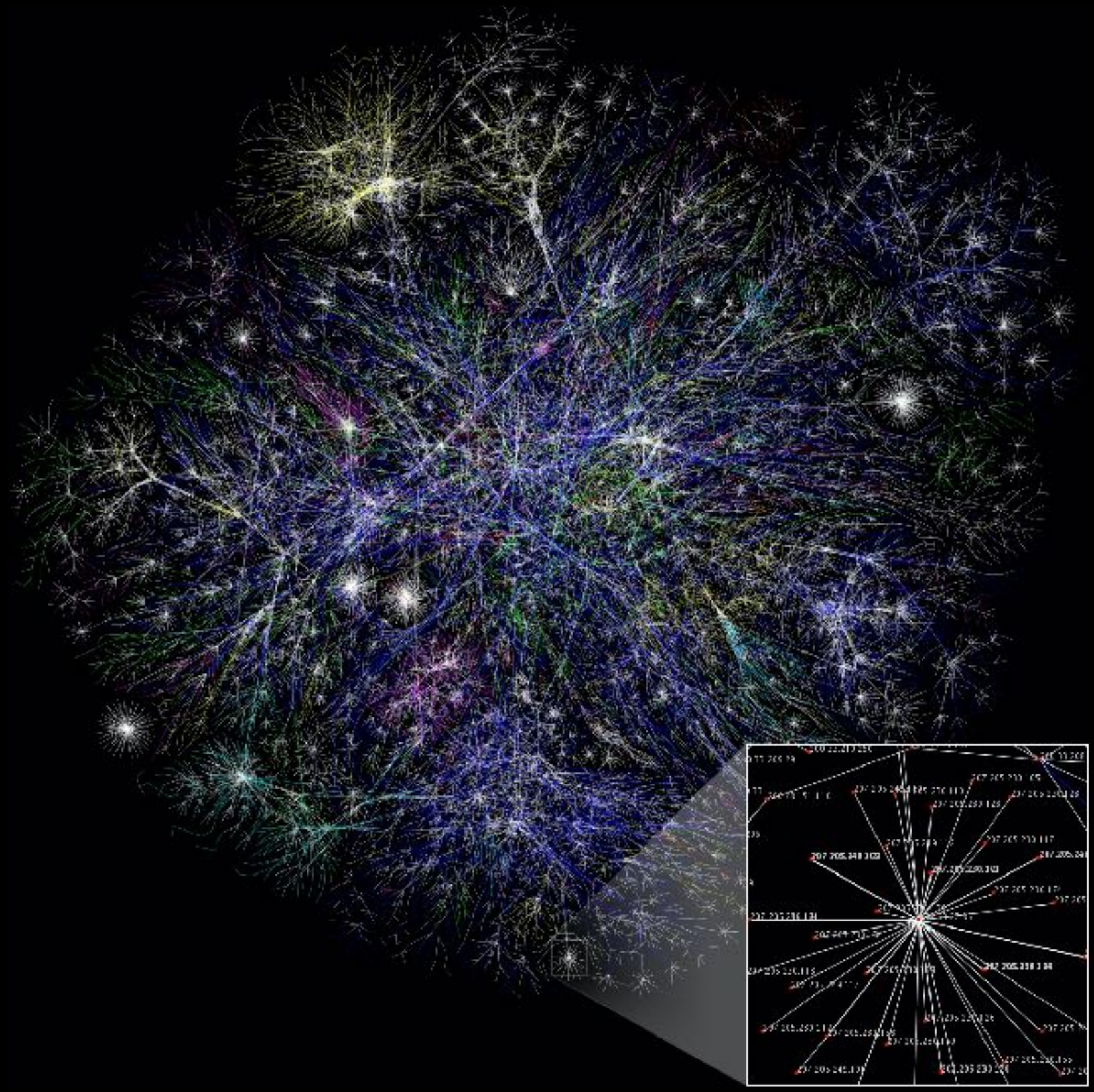
Testing hypotheses?

Effectiveness

Once you've decided upon the task

Did it do the job?

Visualisations can be beautiful - but not effective!



By The Opte Project [CC BY 2.5 (https://creativecommons.org/licenses/by/2.5)], via Wikimedia Commons

Resource limitation

Resource limitation

Computational limitations

Resource limitation

Computational limitations

e.g. How quickly can a visualisation go through a data set?

Resource limitation

Computational limitations

e.g. How quickly can a visualisation go through a data set?

Human limitations

Resource limitation

Computational limitations

e.g. How quickly can a visualisation go through a data set?

Human limitations

Human memory for items not visible is poor

Resource limitation

Computational limitations

e.g. How quickly can a visualisation go through a data set?

Human limitations

Human memory for items not visible is poor

“Change Blindness” - focus on one thing, ignore others

Resource limitation

Computational limitations

e.g. How quickly can a visualisation go through a data set?

Human limitations

Human memory for items not visible is poor

“Change Blindness” - focus on one thing, ignore others

Information Density

Information Density

Information Density

The information density of a single image is a measure of the amount of information encoded versus the amount of unused space.

Information Density

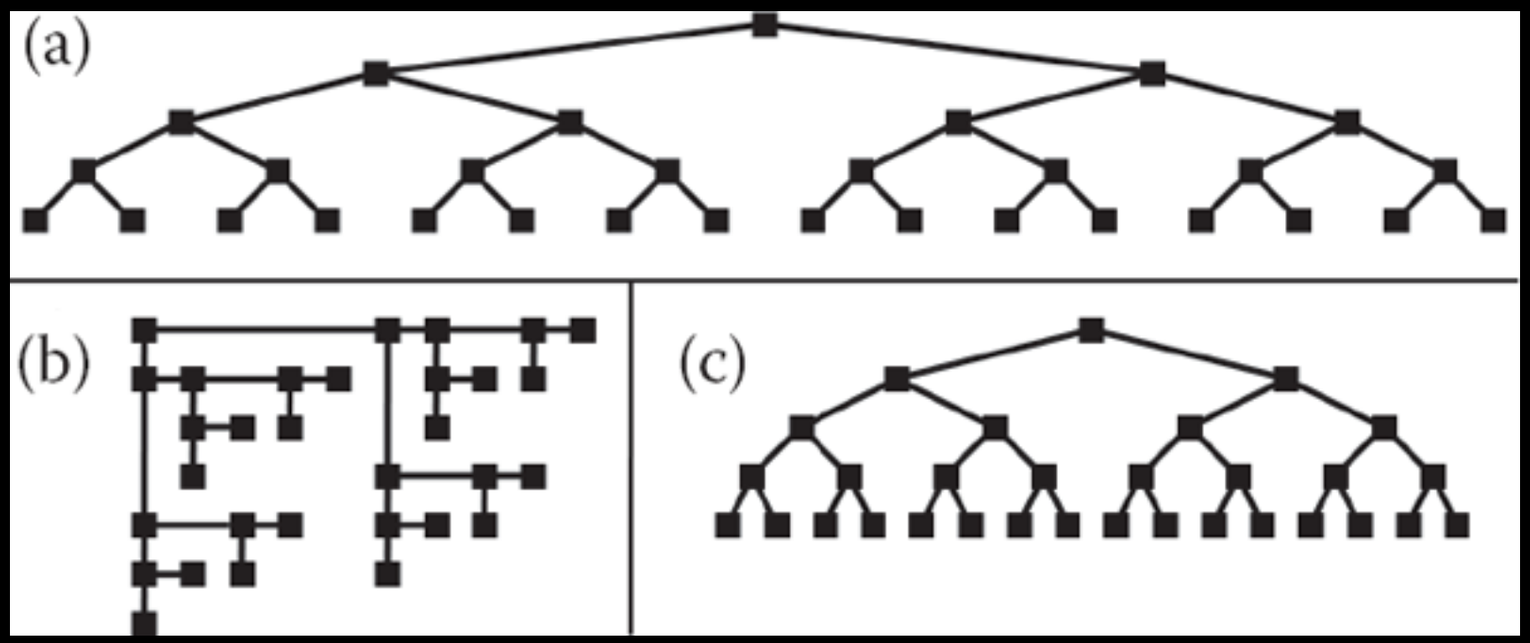
The information density of a single image is a measure of the amount of information encoded versus the amount of unused space.

Information Density also called Graphic Density or Data-Ink Ratio

Information Density

The information density of a single image is a measure of the amount of information encoded versus the amount of unused space.

Information Density also called Graphic Density or Data-Ink Ratio

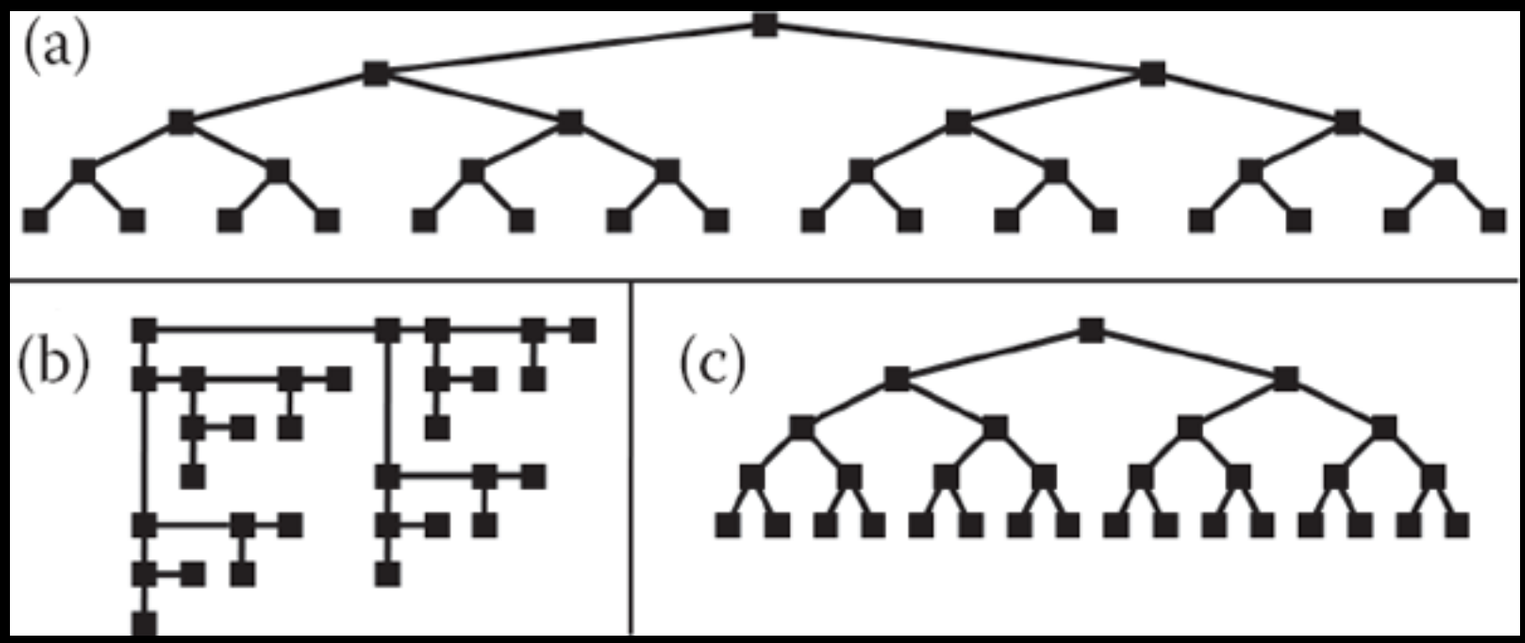


Information Density

The information density of a single image is a measure of the amount of information encoded versus the amount of unused space.

Information Density also called Graphic Density or Data-Ink Ratio

Low info. Density



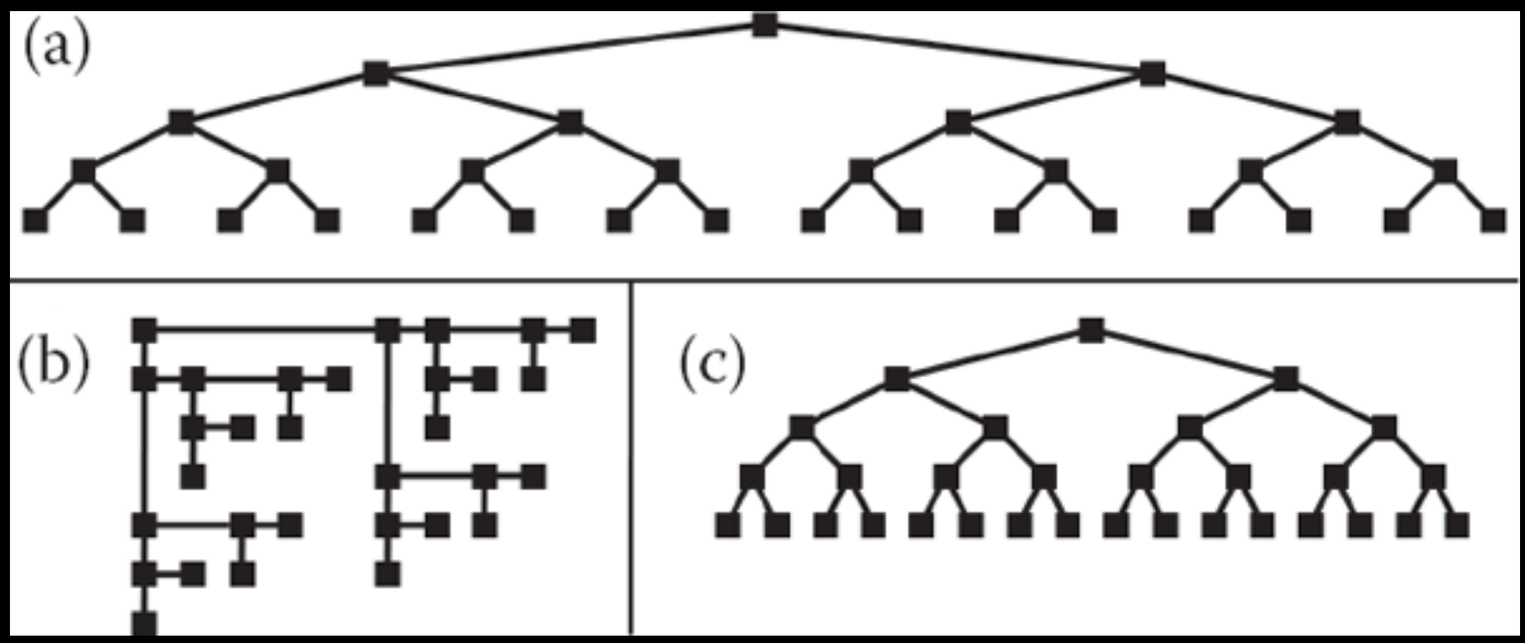
Information Density

The information density of a single image is a measure of the amount of information encoded versus the amount of unused space.

Information Density also called Graphic Density or Data-Ink Ratio

Low info. Density

High Info. Density (but can't tell tree order)



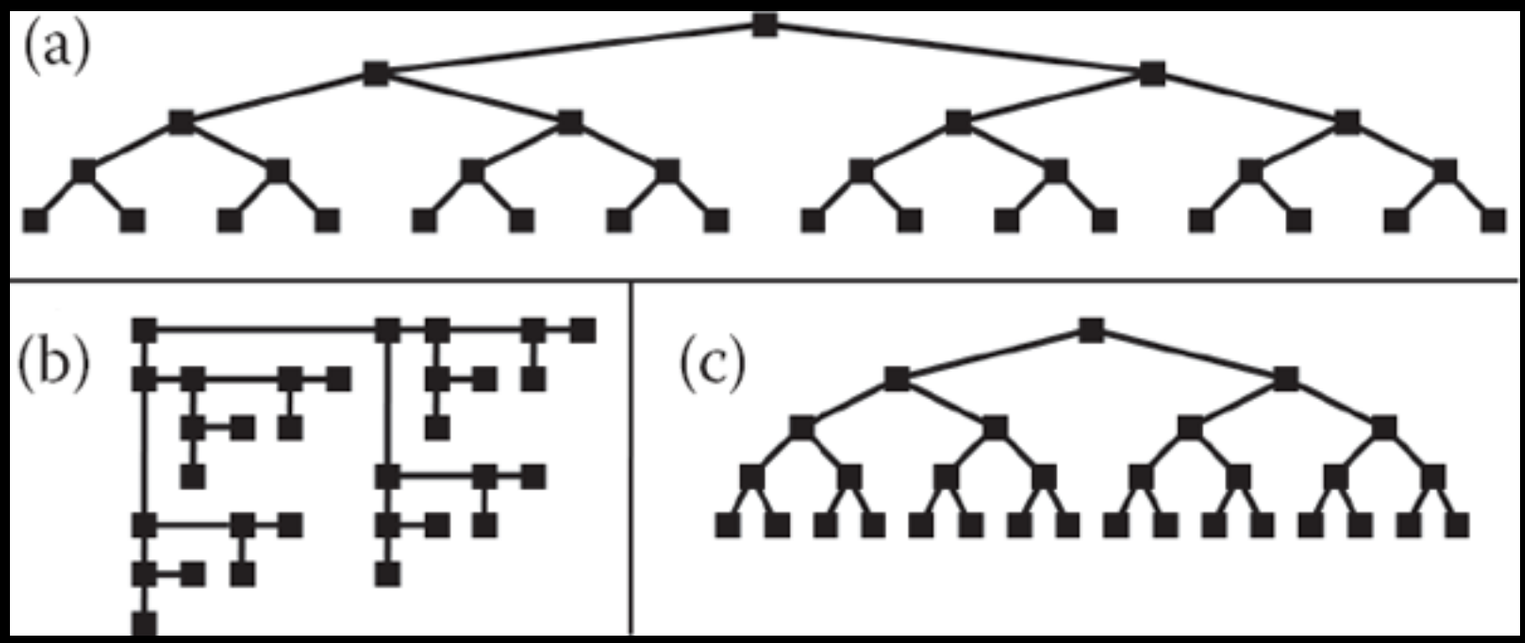
Information Density

The information density of a single image is a measure of the amount of information encoded versus the amount of unused space.

Information Density also called Graphic Density or Data-Ink Ratio

Low info. Density

High Info. Density (but can't tell tree order)



High Info. Density and can Tell tree order

Histograms and Distributions

So far…

Scatterplots - pairs of attributes

How does a single attribute behave?

Example

Life expectancy in GapMinder data set

Example

Life expectancy in GapMinder data set

Varies between countries

1967 Average Life Expectancy in Morocco, Bangladesh = 50.34, 43.45

Example

Life expectancy in GapMinder data set

Varies between countries

1967 Average Life Expectancy in Morocco, Bangladesh = 50.34, 43.45

Varies over time

Average life expectancy in Namibia in 1952, 2007 = 41.73, 52.90

Aside - life expectancy for country itself an estimate

Varies according to

Year of birth

Socioeconomic group (wealthy/poor)

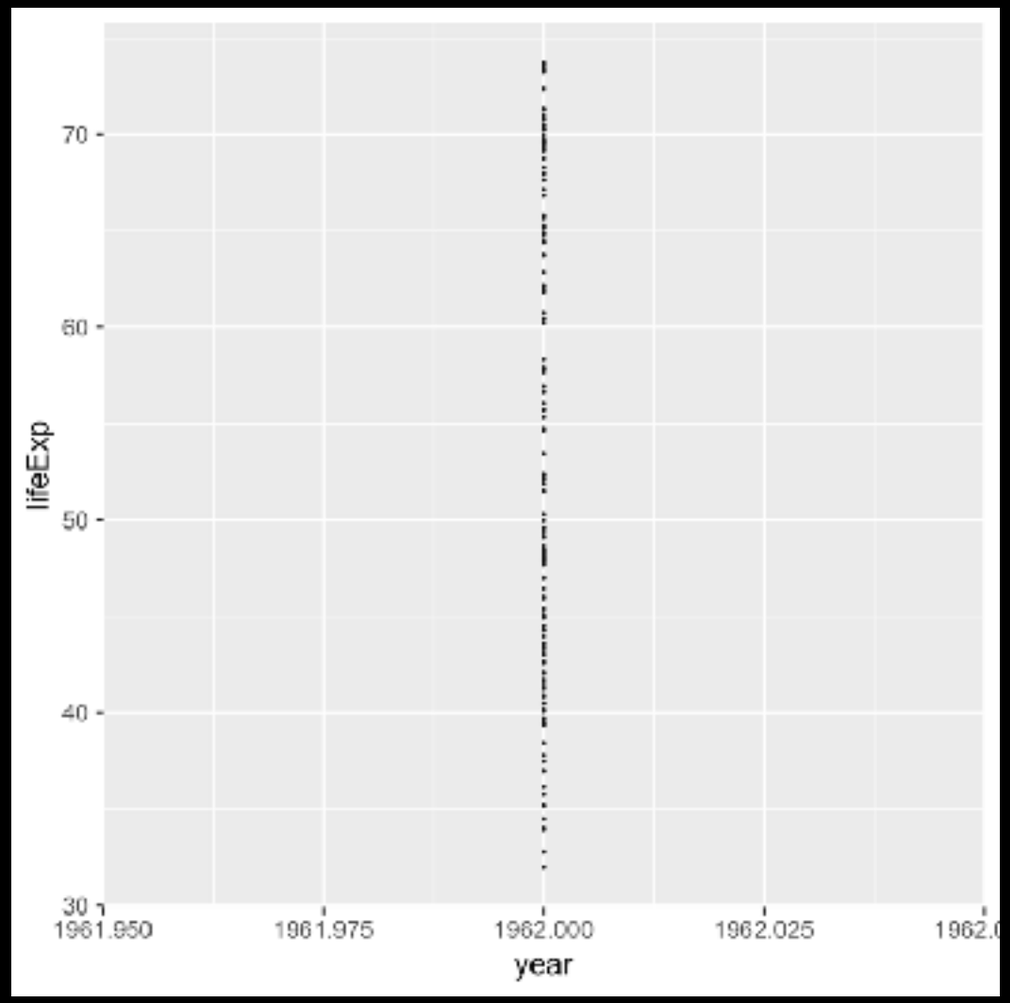
Geographical location

First attempt - just plot points

Example - Life Expectancy in 1962 for all countries

First attempt - just plot points

Example - Life Expectancy in 1962 for all countries

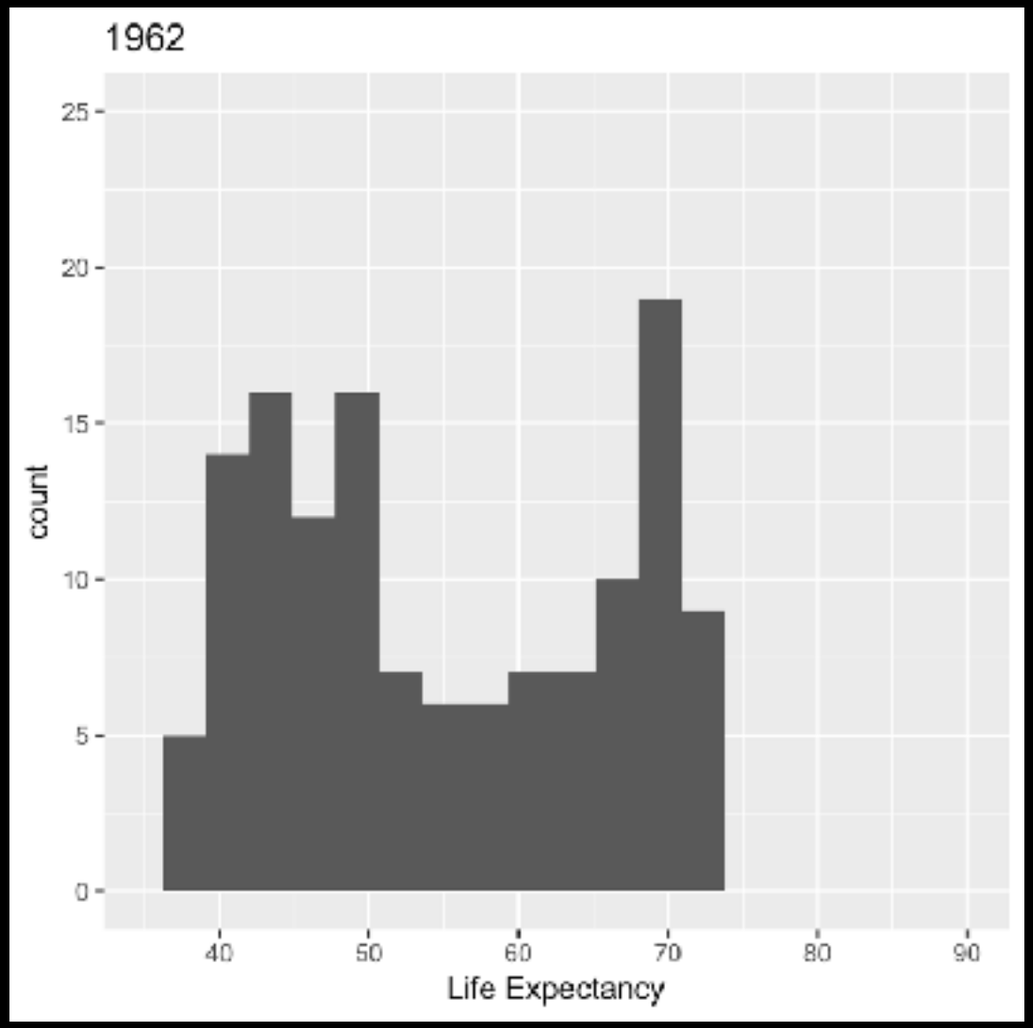


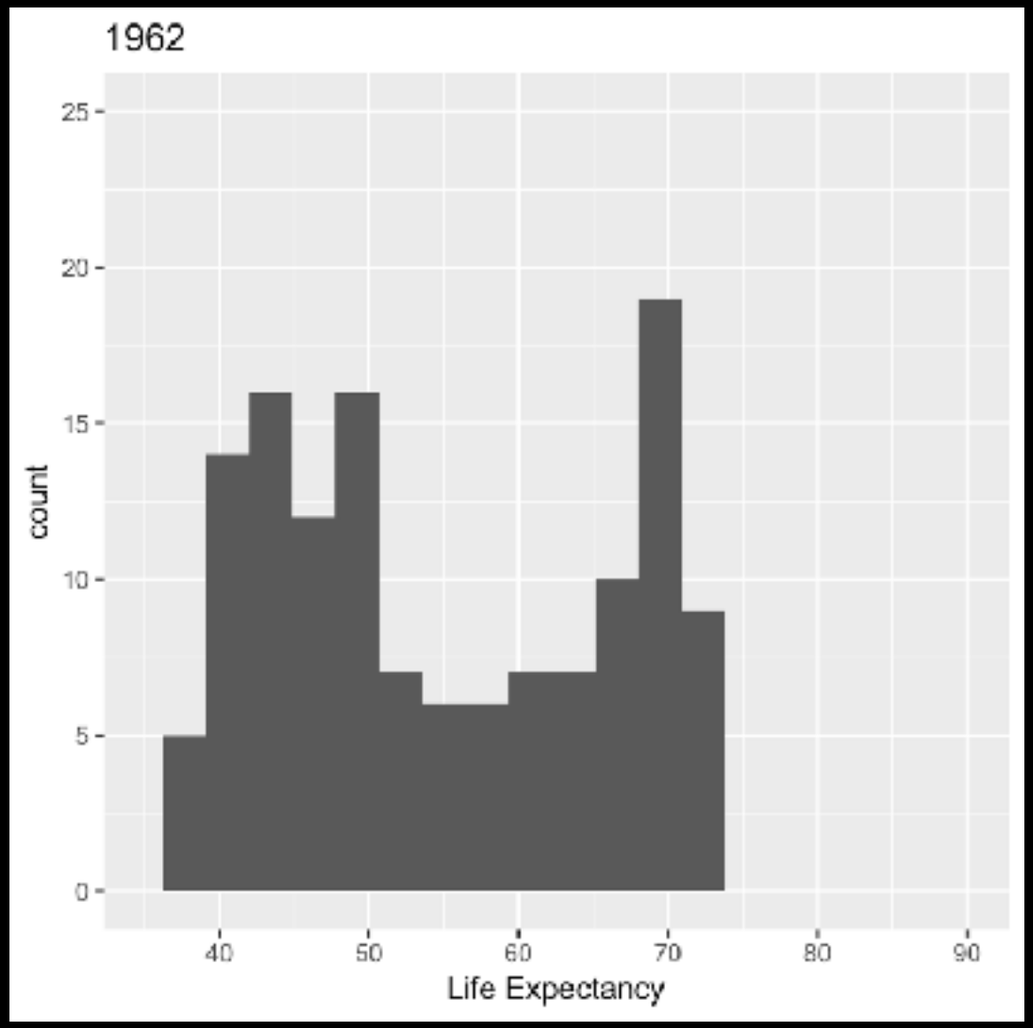
Histogram

Means of converting Ordinal to Categorical data

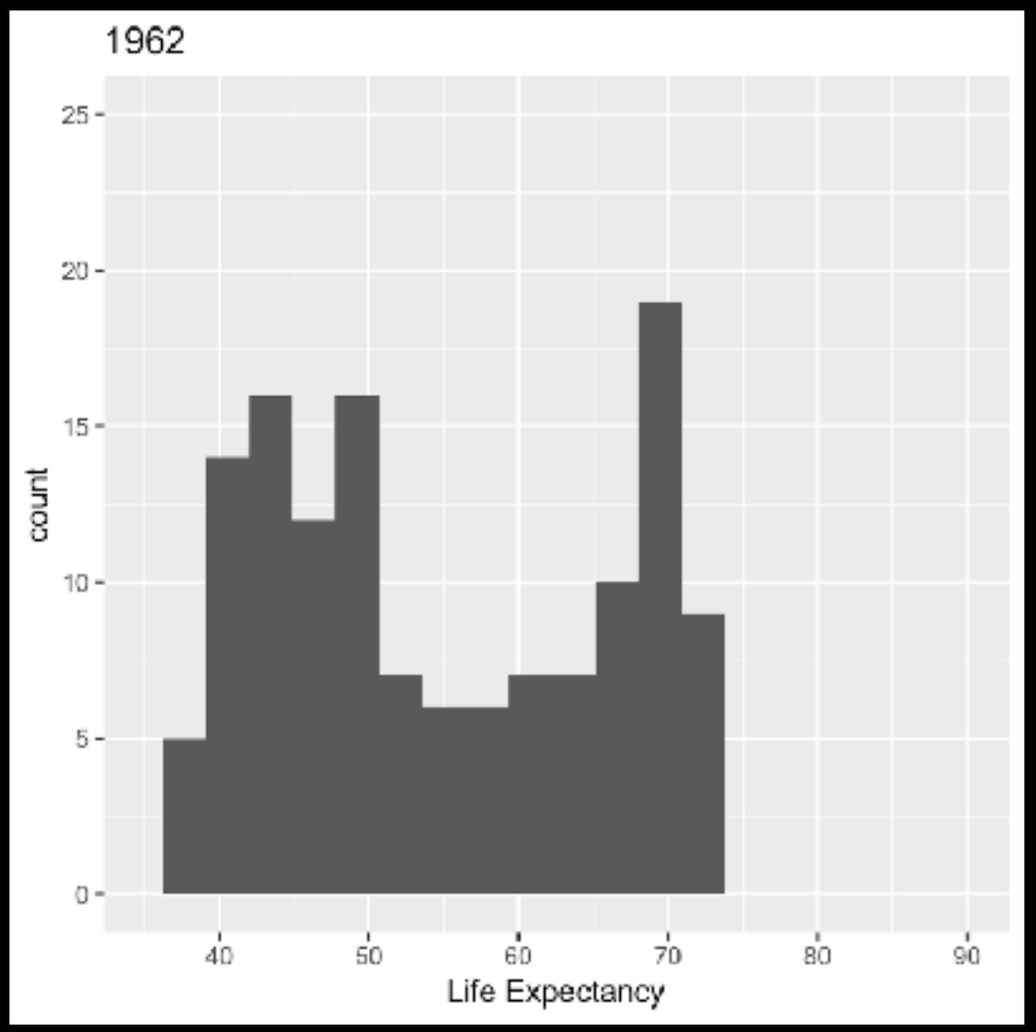
Create ‘bins’ -
how many countries have an Life Expectancy of 40-45 etc.

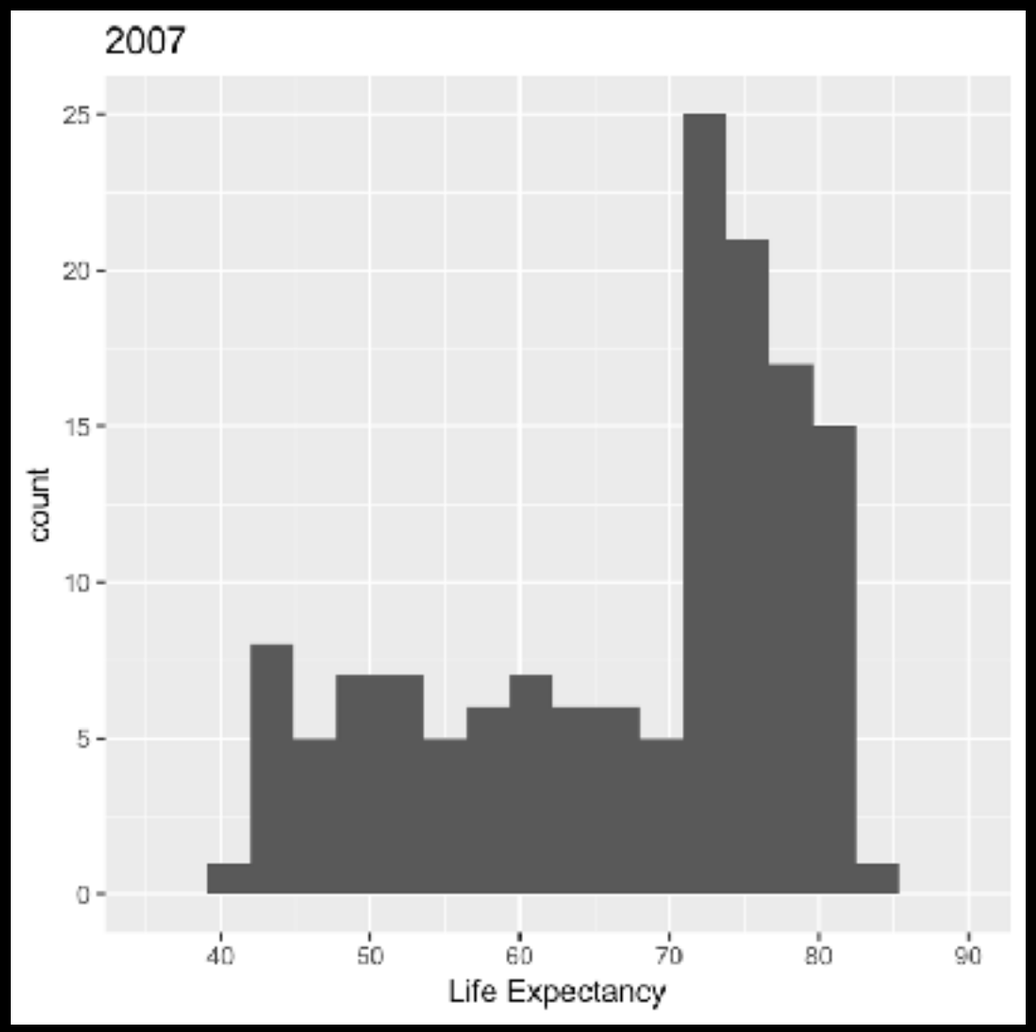
Create bar plot of counts


1962
25
20
15
10
5
0
count
40
50
60
70
80
90
Life Expectancy

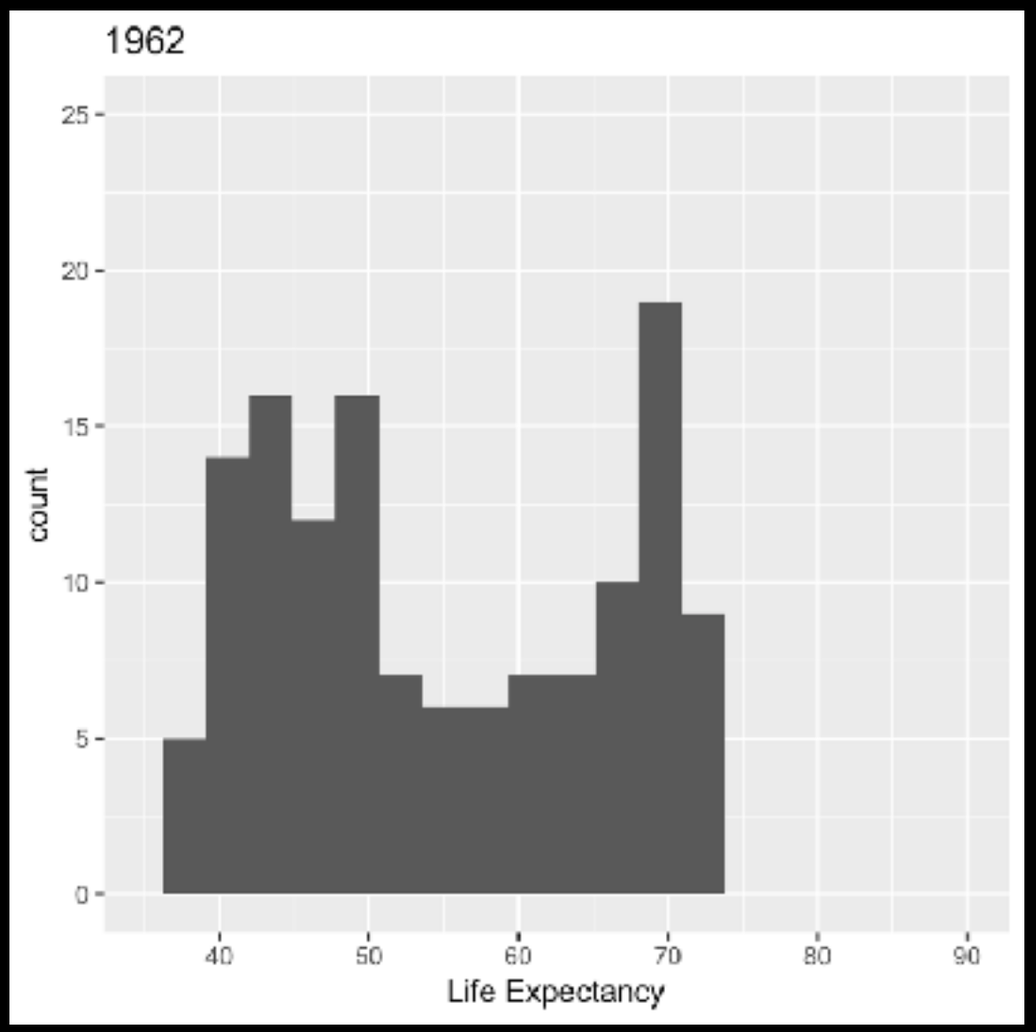


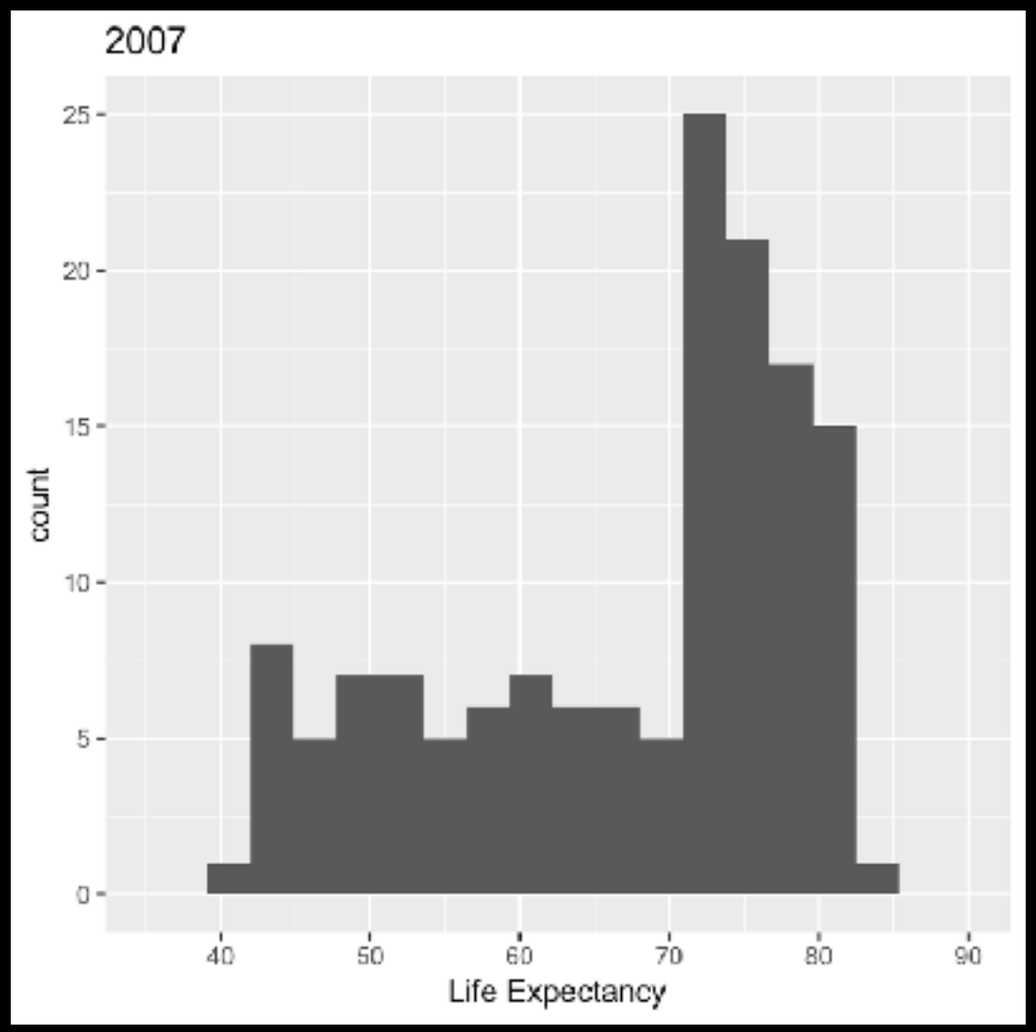
What are seeing here?





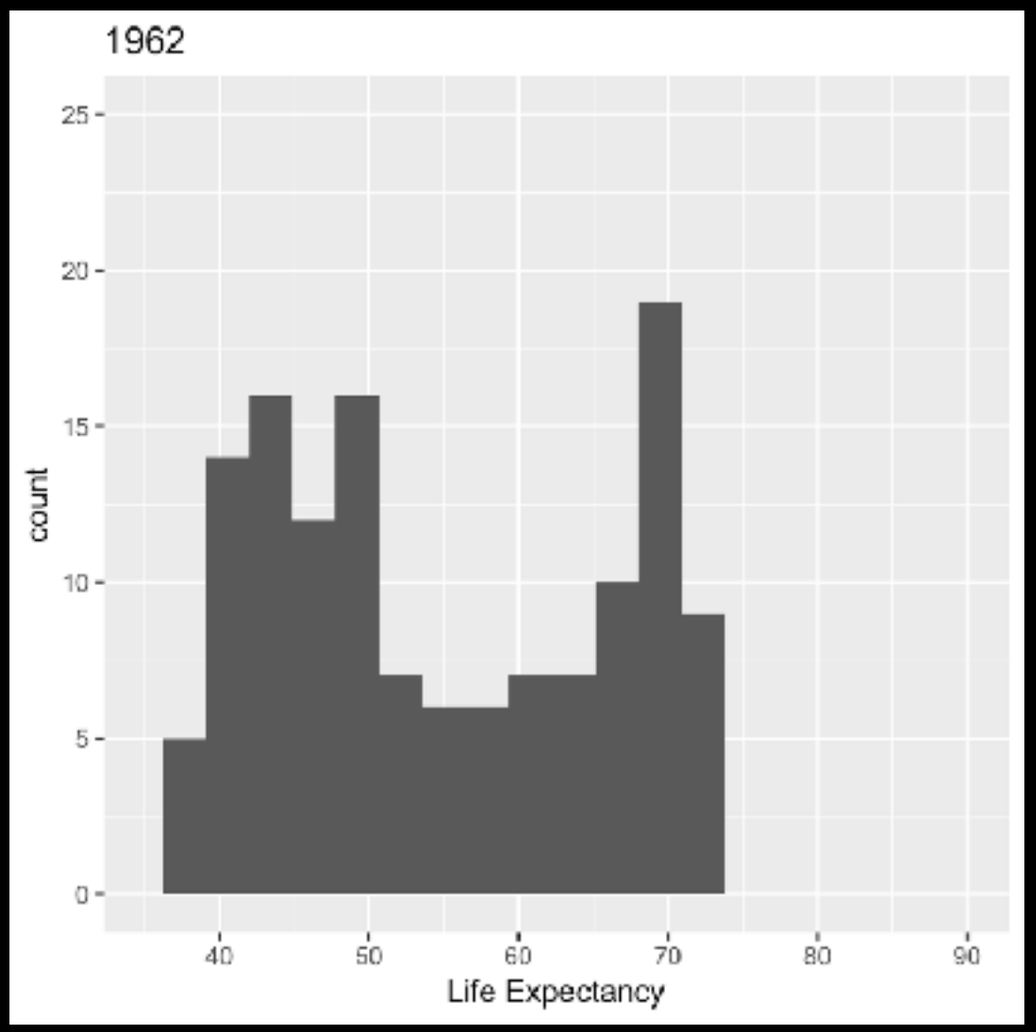
What are seeing here?

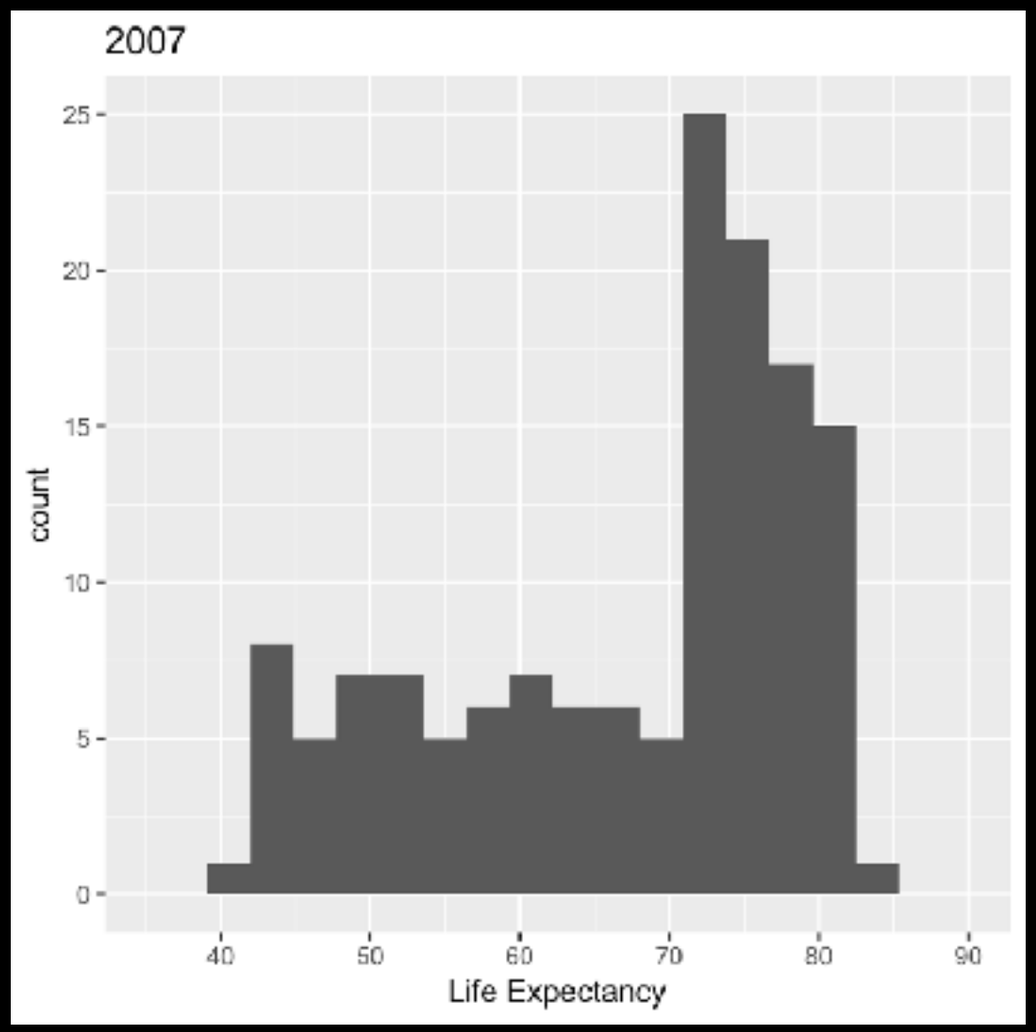




What are seeing here?

What's different?





What are seeing here?

What's different?

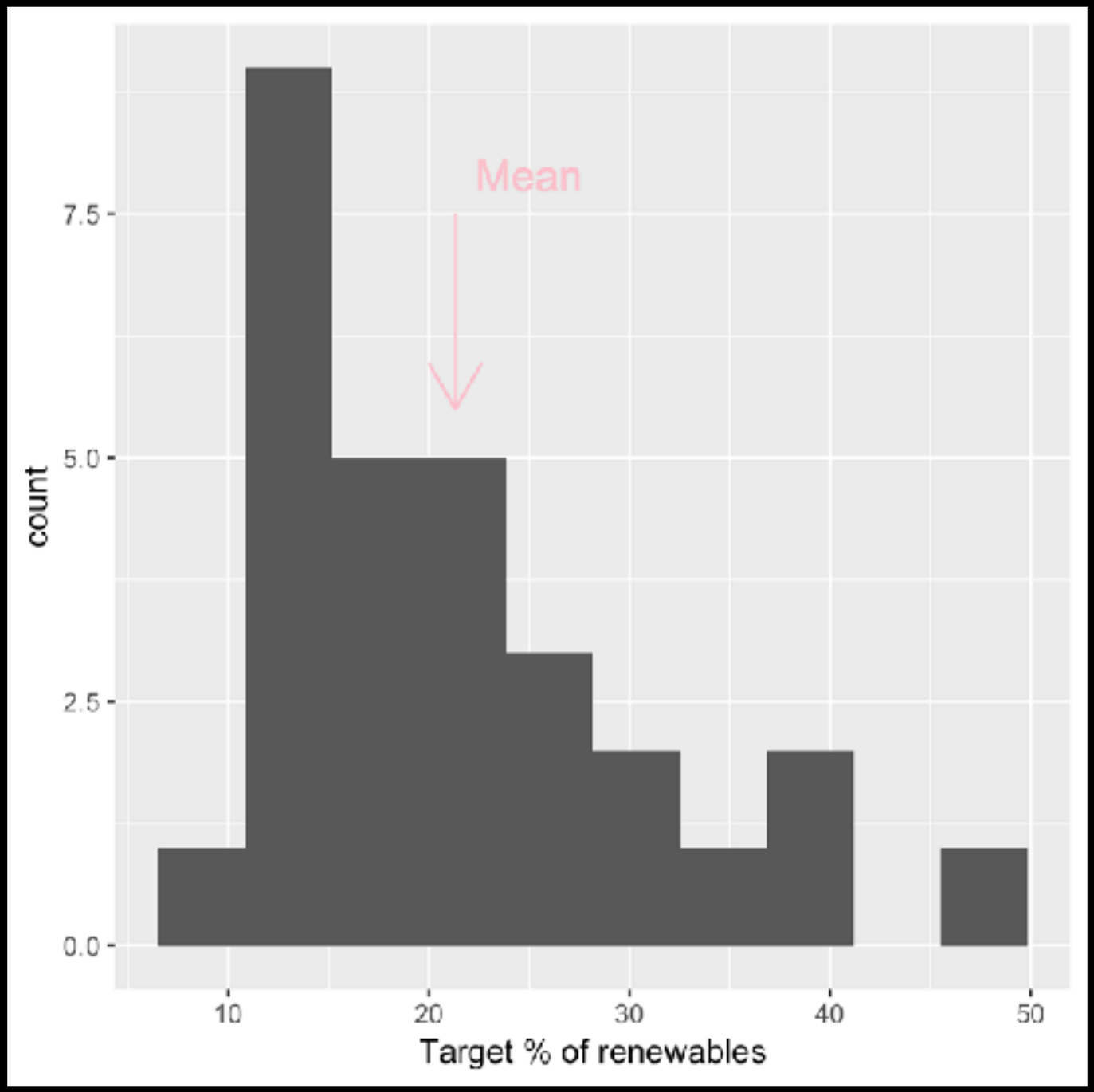
What did I do to make this comparison?

Number of bins

Too many - just see noise

Too few - don’t see any features

What about this ?

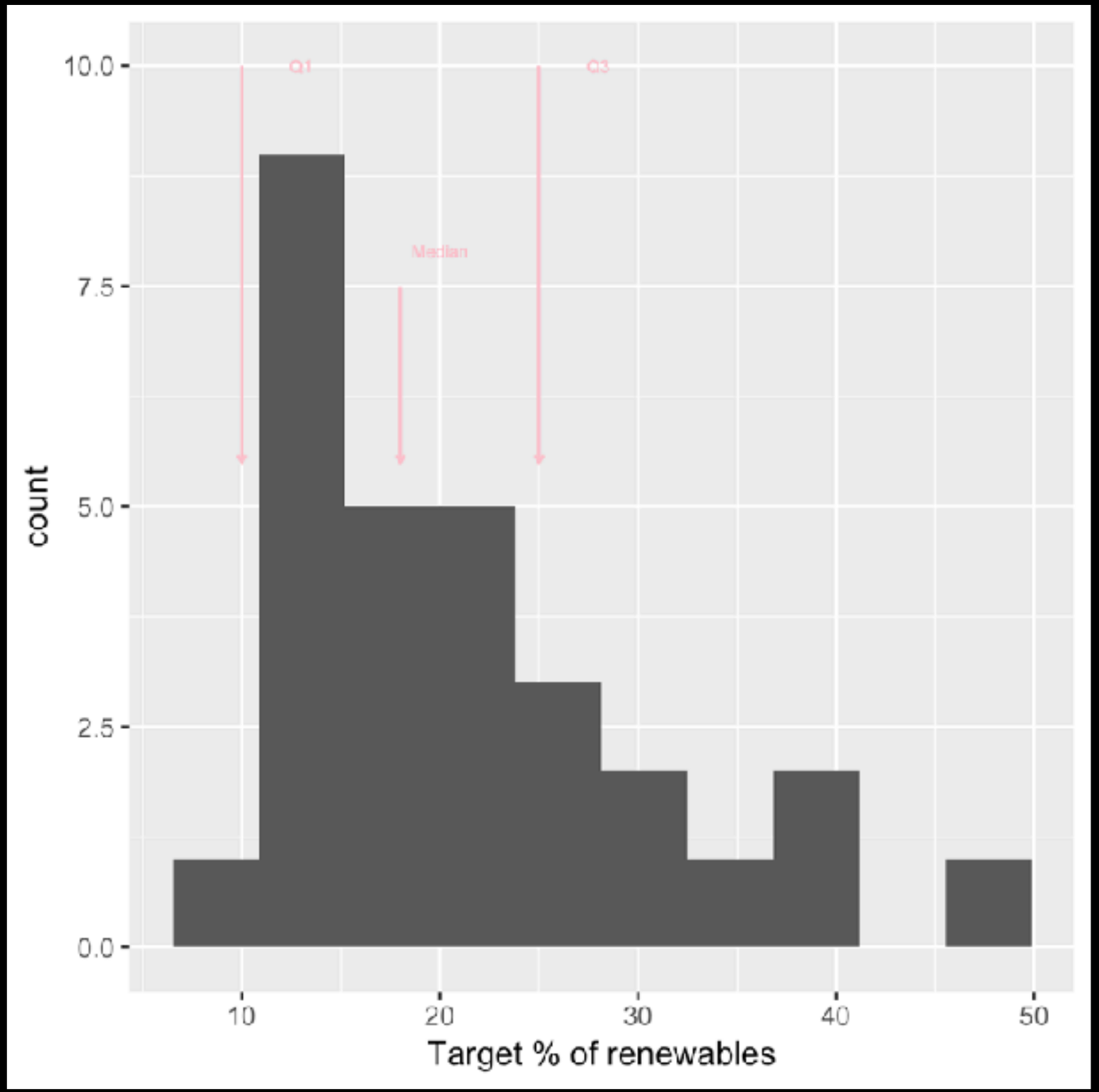


Better to think of quartiles

1st quartile (Q1) - value of y which is greater than 25% of the y_i

Median - value of y which is greater than 50% of the y_i

3rd quartile (Q3) - value of y which is greater than 75% of the y_i


Q1
Median
Q3
10.0
7.5
5.0
2.5
0.0
count
10
20
30
40
50
Target % of renewables

What if we want to compare distributions?

Life expectancy of countries between years?

What if we want to compare distributions?

Life expectancy of countries between years?

Could do a facet plot of histograms

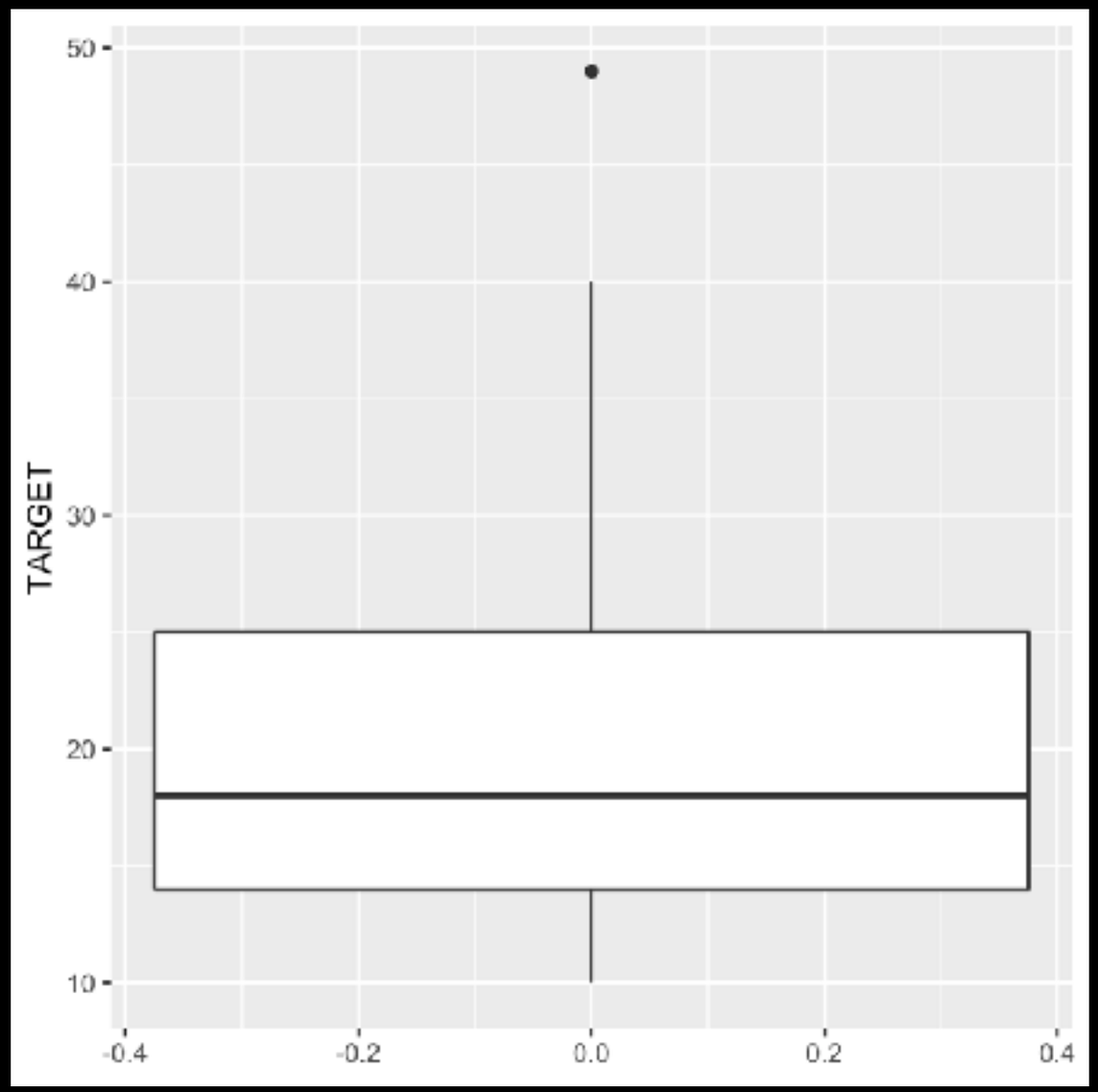
What if we want to compare distributions?

Life expectancy of countries between years?

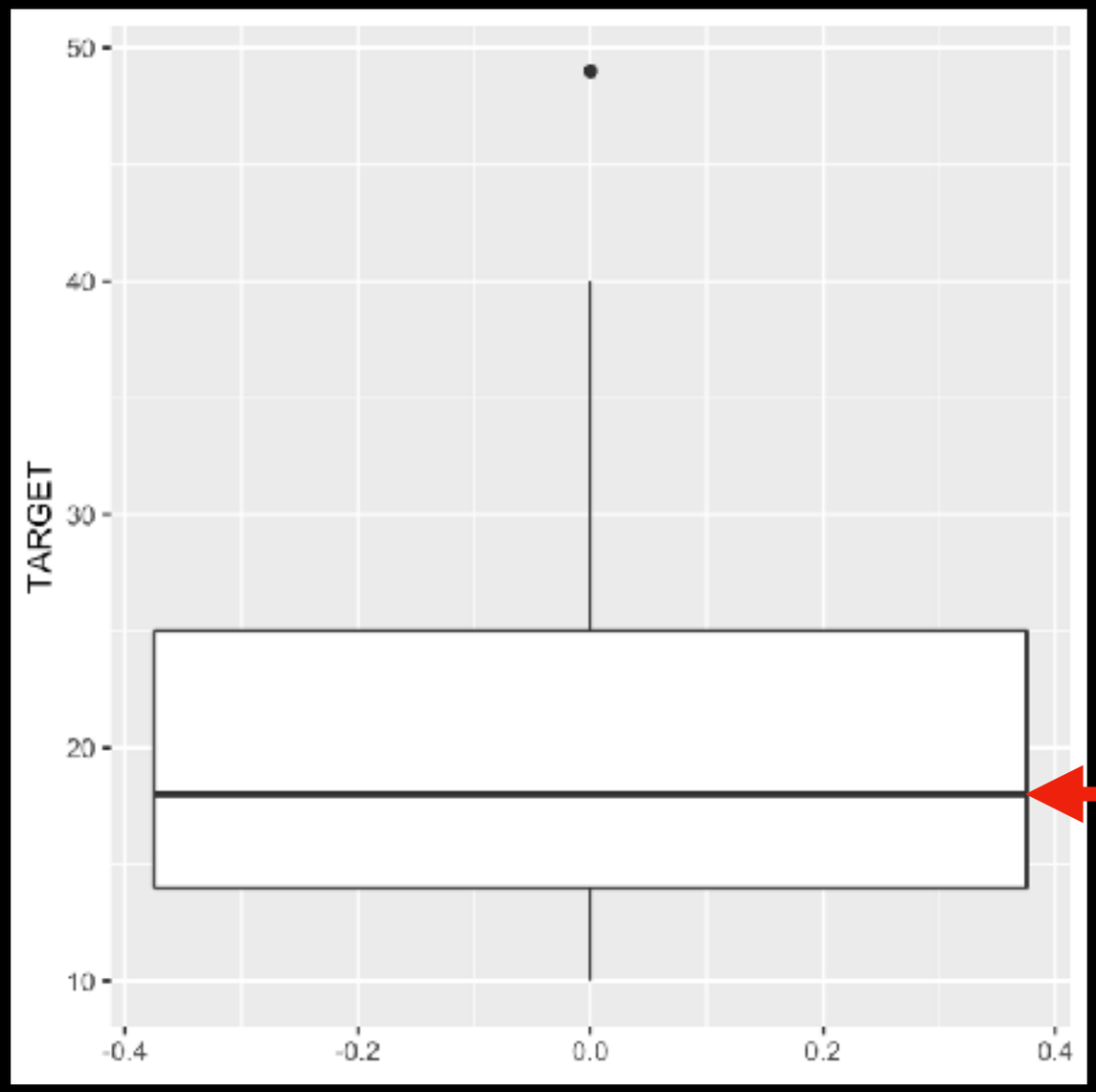
Could do a facet plot of histograms

But can also other comparisons

Box plot of renewables target

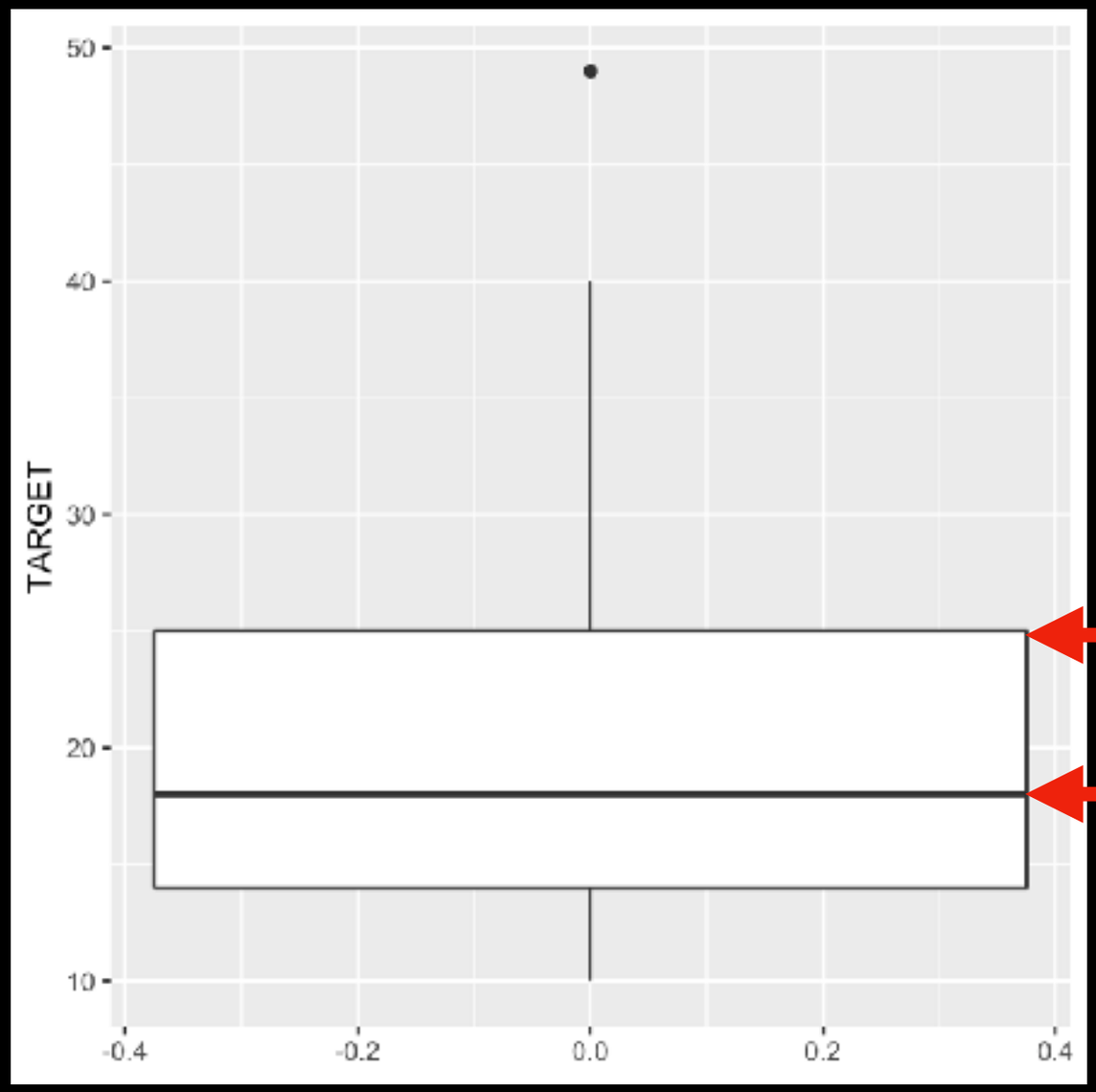


Box plot of renewables target



Median

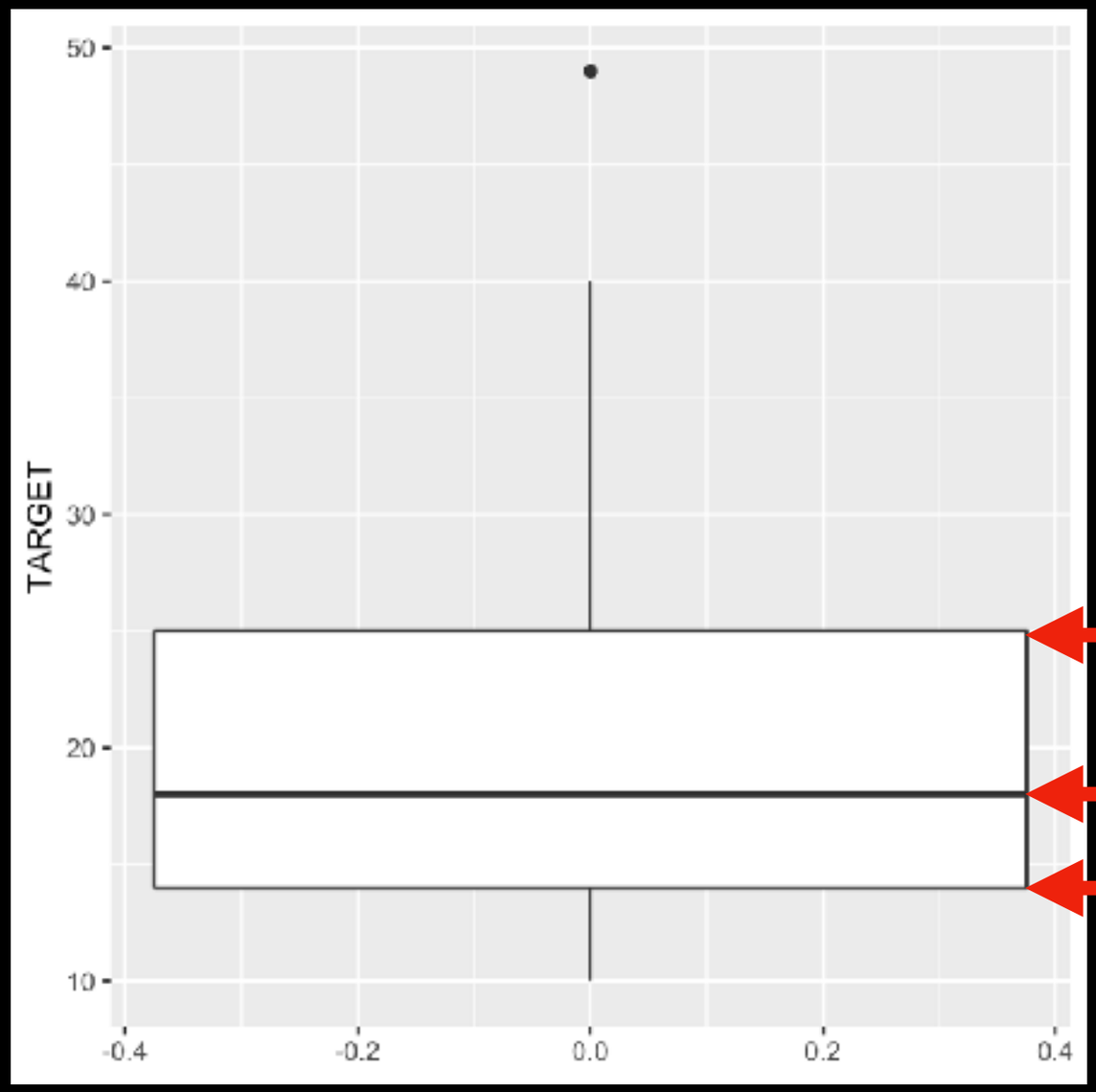
Box plot of renewables target



Q3

Median

Box plot of renewables target

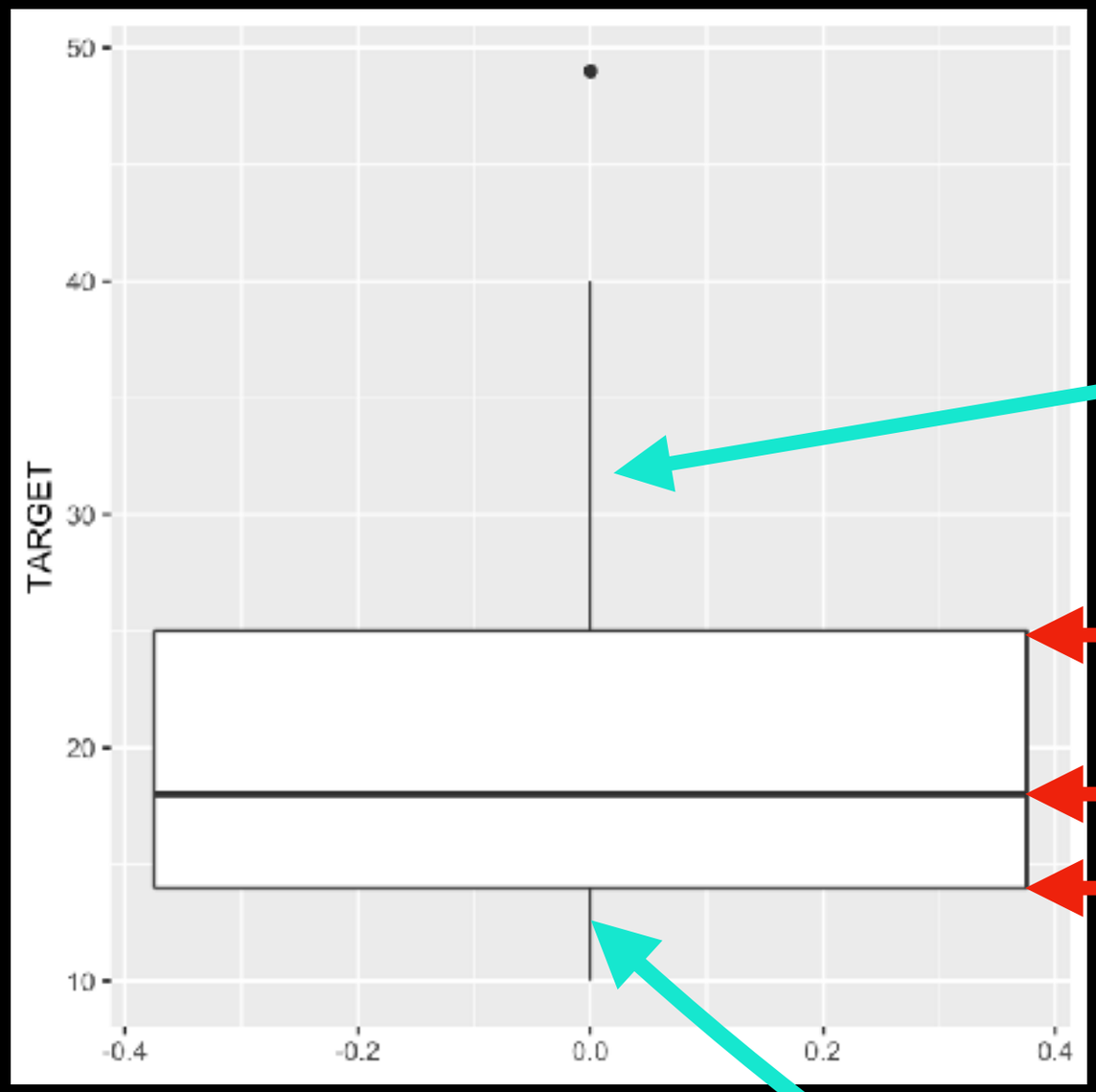


Q3

Median

Q1

Box plot of renewables target



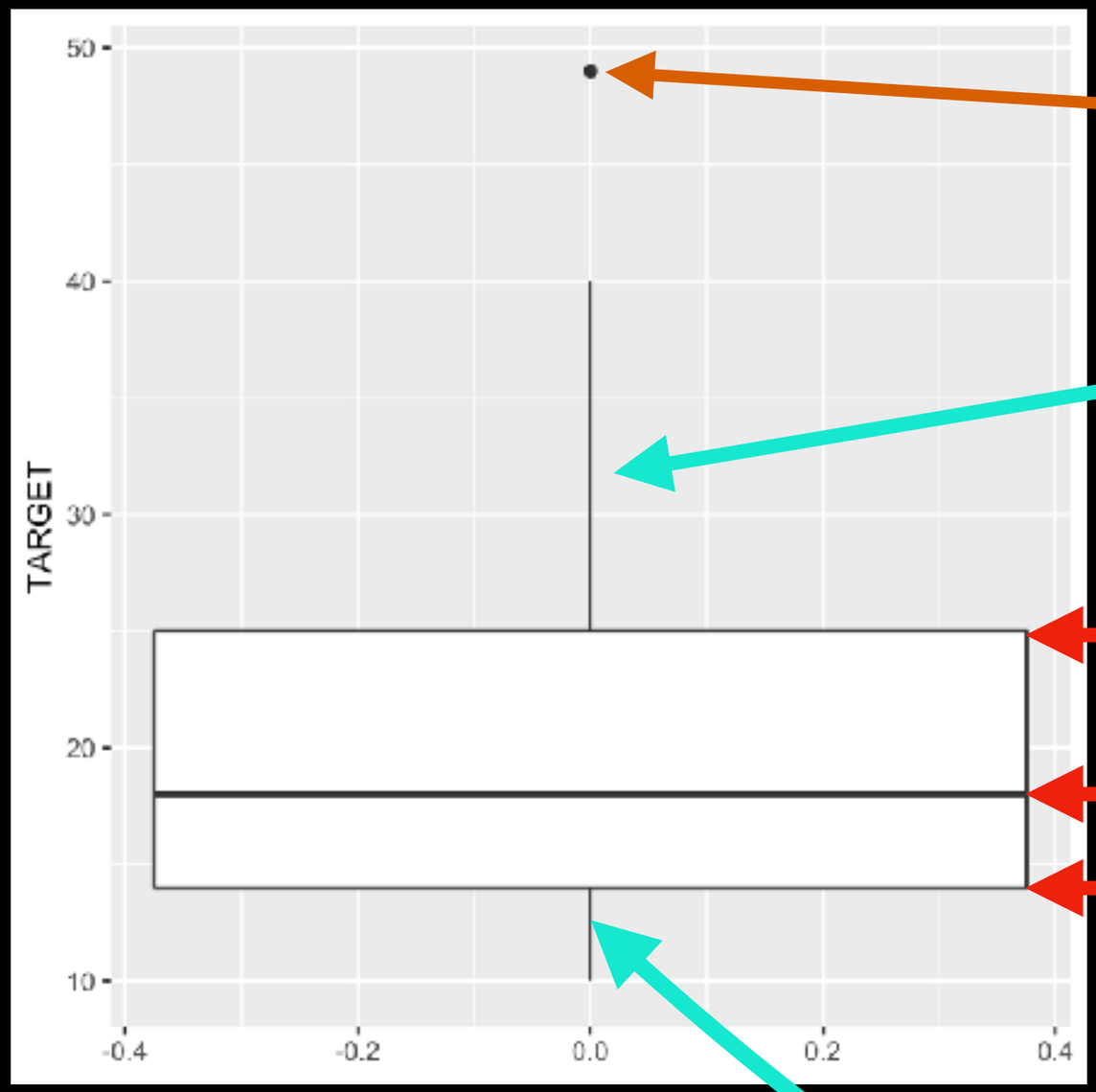
Whiskers

Q3

Median

Q1

Box plot of renewables target



Outlier

Whiskers

Q3

Median

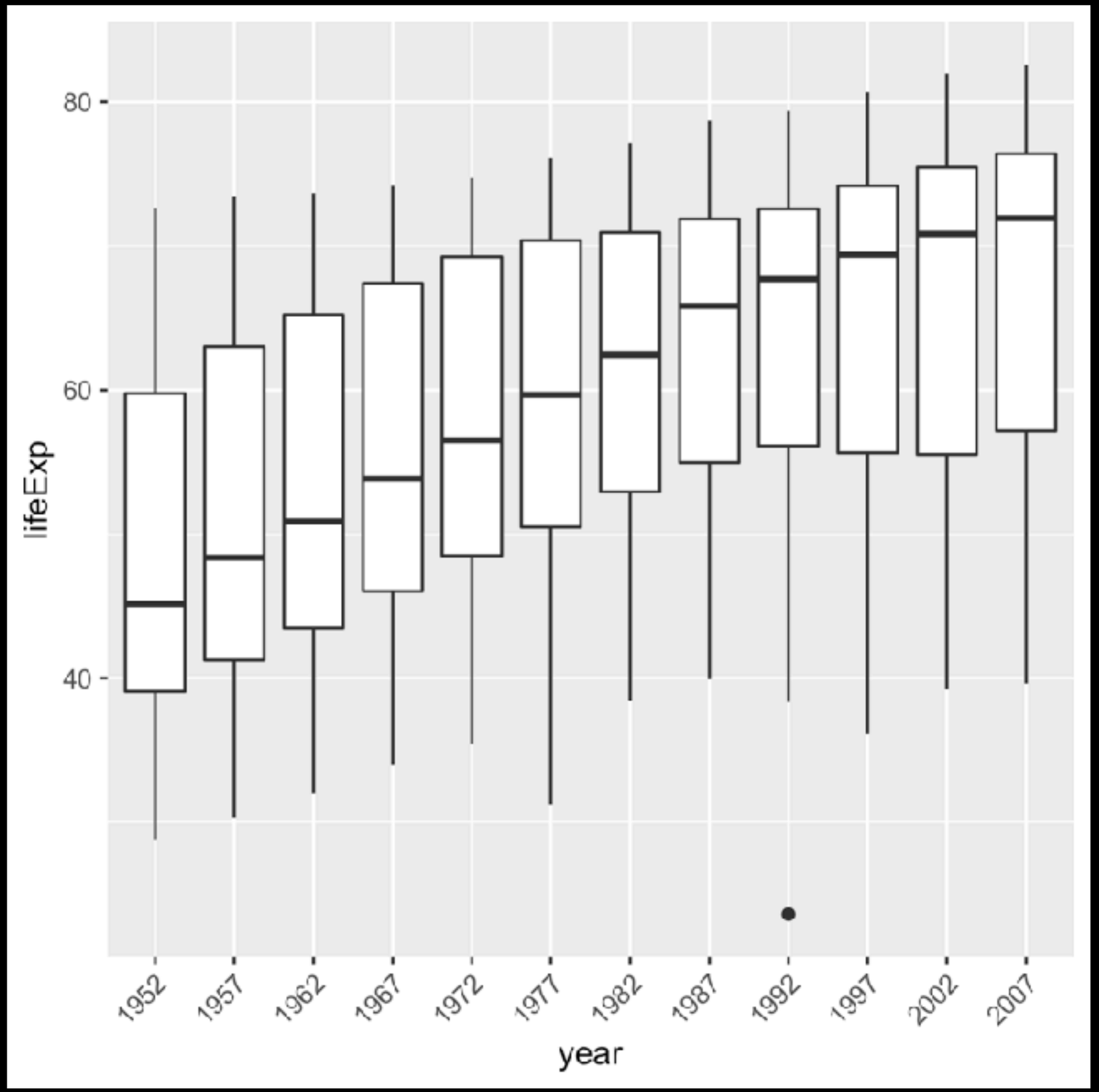
Q1

Whisker length

No bigger than 1.5 x IQR

IQR = Q3 - Q1

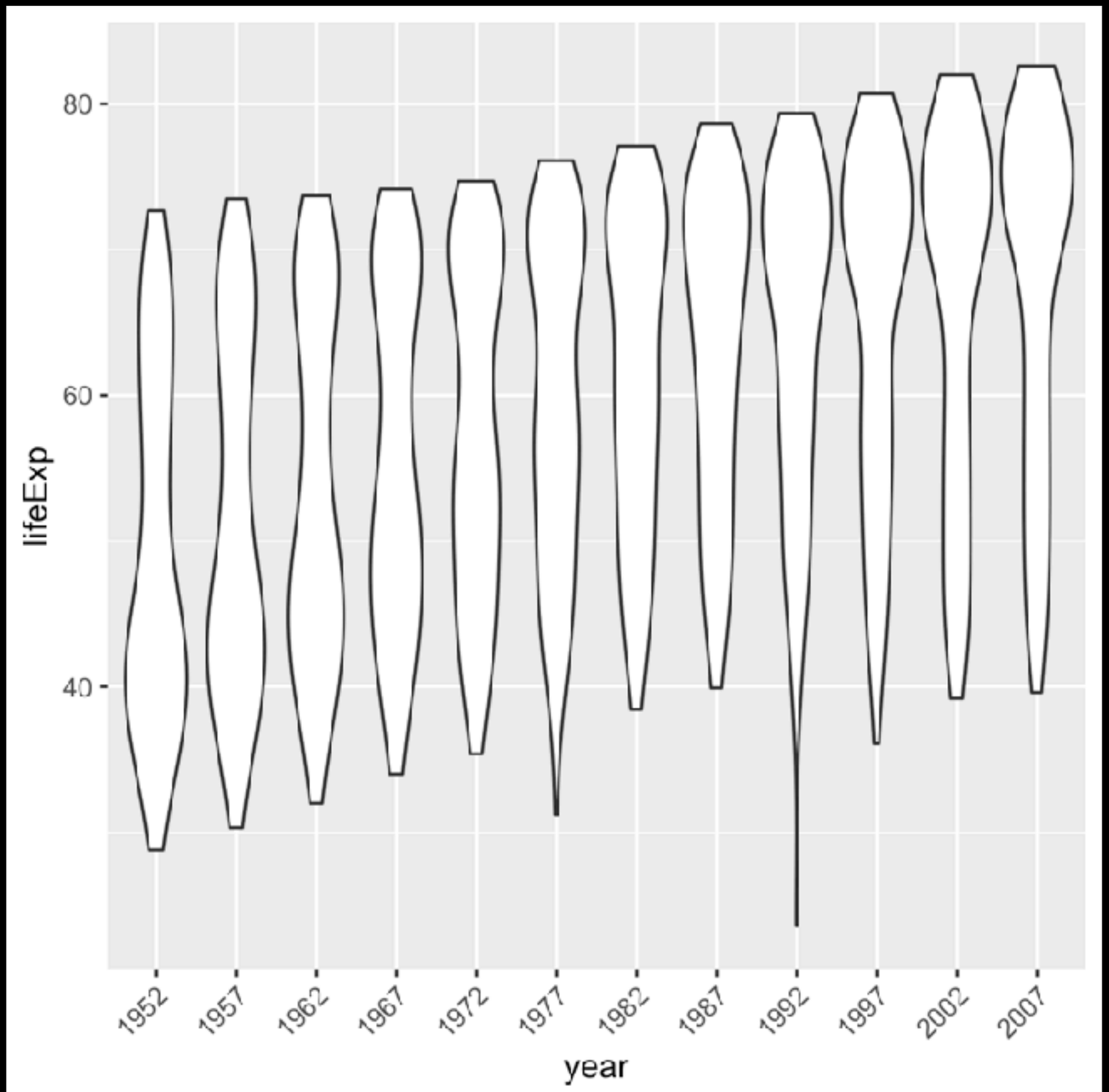
If any data is greater than maximum whisker length then plotted as point


80
60
40
lifeExp
1952
1957
1962
1967
1972
1977
1982
1987
1992
1997
2002
2007
year

But Life Expectancy isn’t unimodal !

Try Violin plots

Estimate distribution - width of shape indicates height of distribution


80
60
40
lifeExp
1952
1957
1962
1967
1972
1977
1982
1987
1992
1997
2002
2007
year

Marks and Visual Channels

Basic building blocks of visualisation

Finite number

Understand their relevance

Marks

Marks

Basic graphical element in an image

Marks

Basic graphical element in an image

Geometric primitives

Marks

Basic graphical element in an image

Geometric primitives

Defined by dimension

Marks

Basic graphical element in an image

Geometric primitives

Defined by dimension

0 Dimension - point

Marks

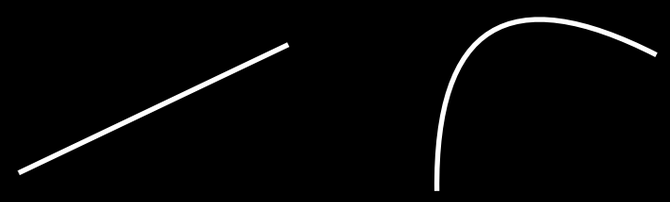
Basic graphical element in an image

Geometric primitives

Defined by dimension

0 Dimension - point

1 Dimension - line



Marks

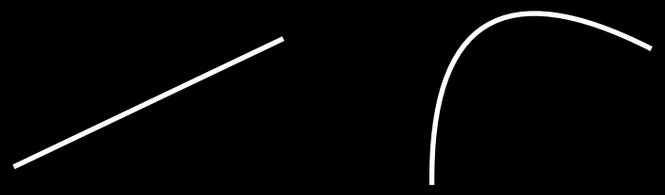
Basic graphical element in an image

Geometric primitives

Defined by dimension

0 Dimension - point

1 Dimension - line



2 Dimensions - area

Can also have three dimensional primitives…..

Can also have three dimensional primitives…..

But we avoid for later reasons

Visual channel

Visual channel

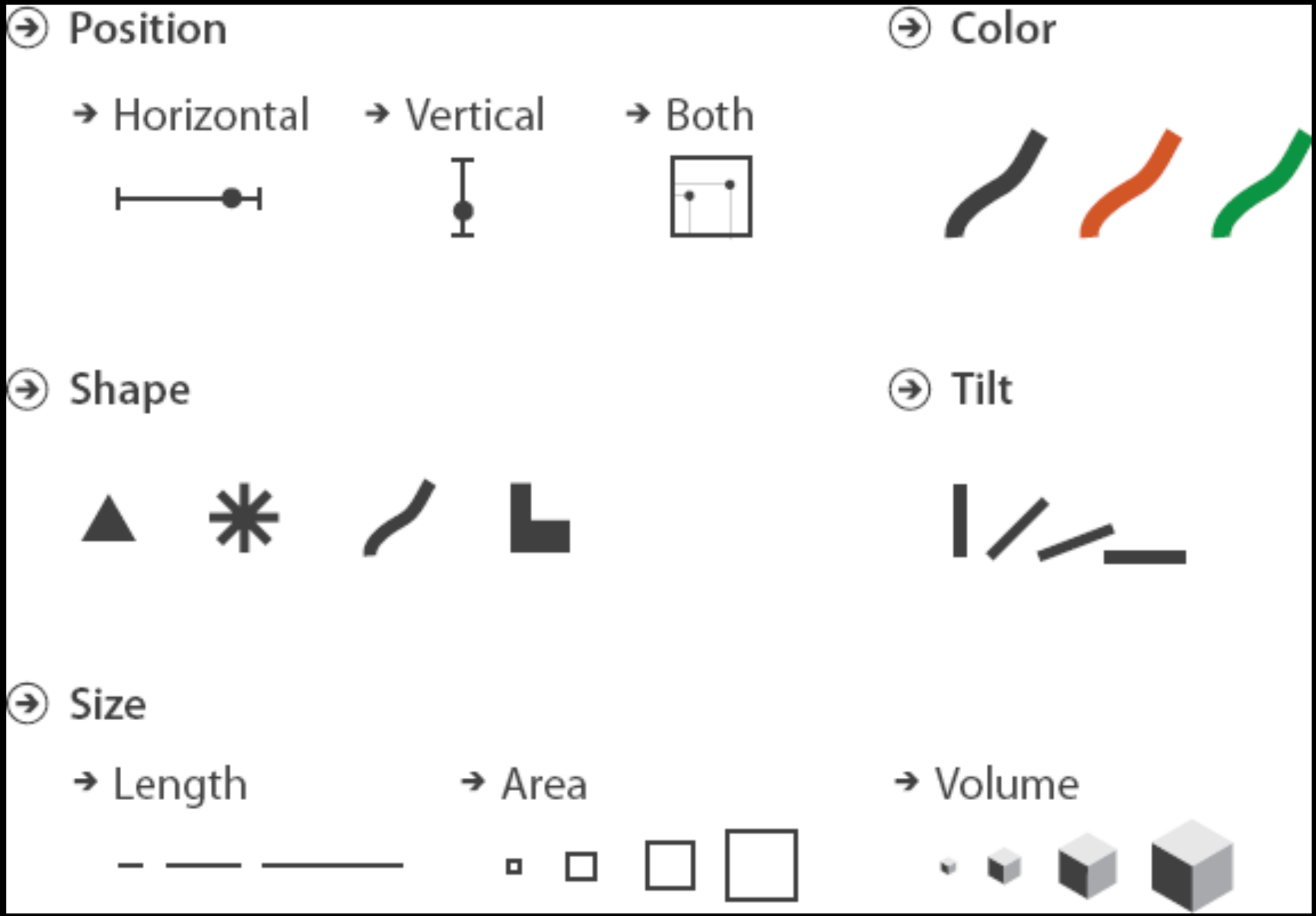
a way to control the appearance of marks, independent of the dimensionality of the geometric primitive.

Visual channel

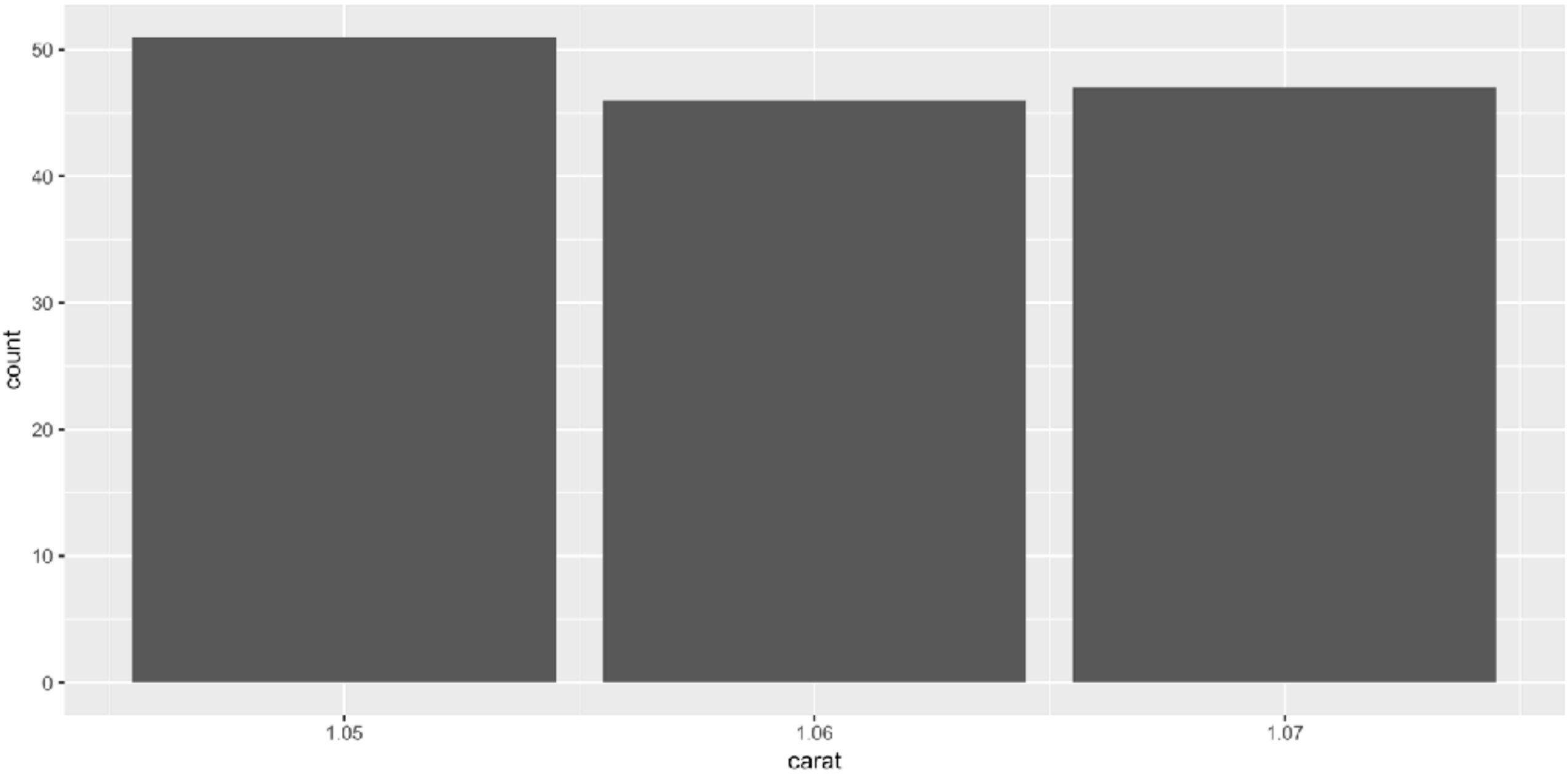
a way to control the appearance of marks, independent of the dimensionality of the geometric primitive.

Many synonyms for this term !

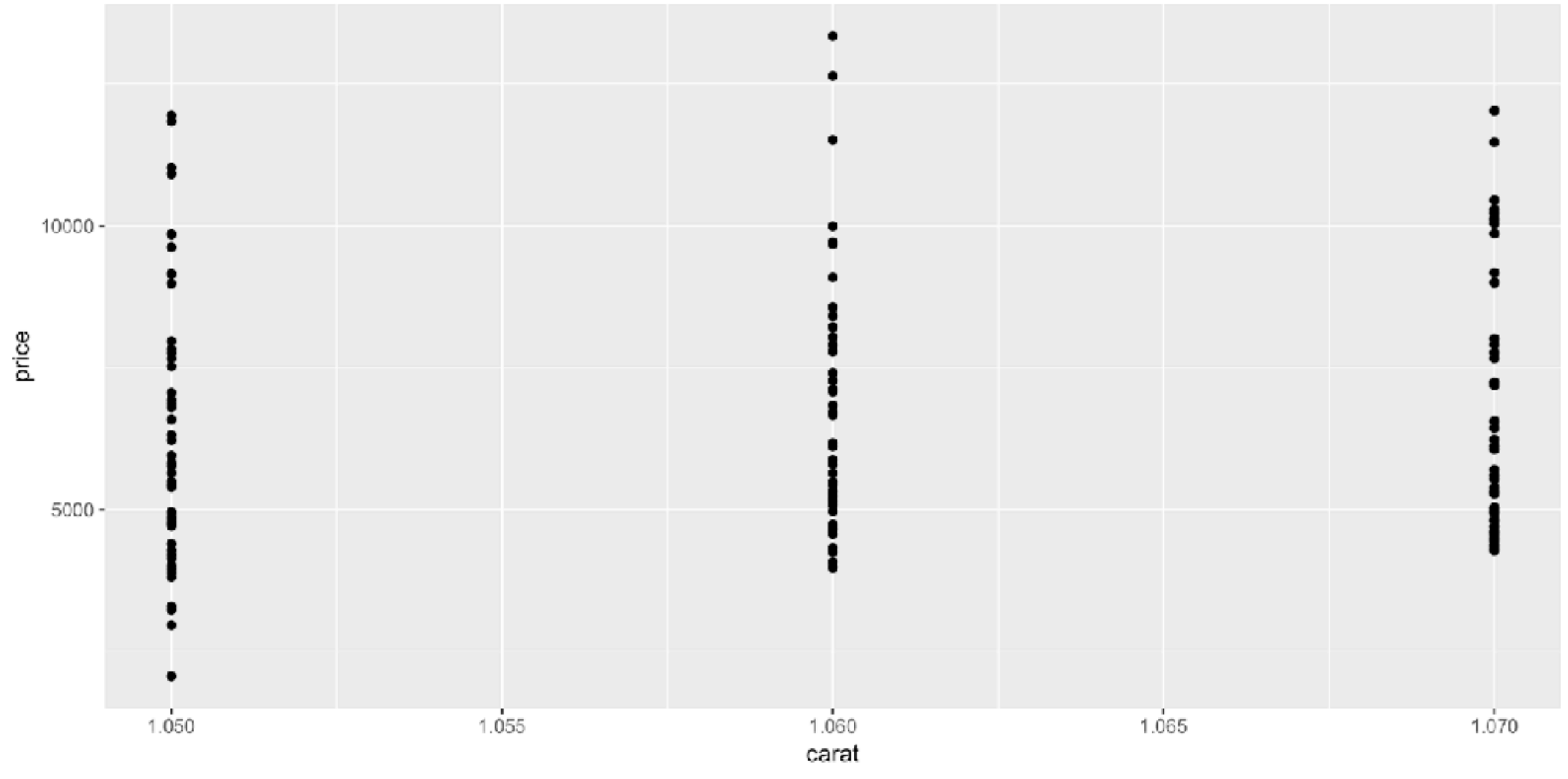
Examples of Visual Channels



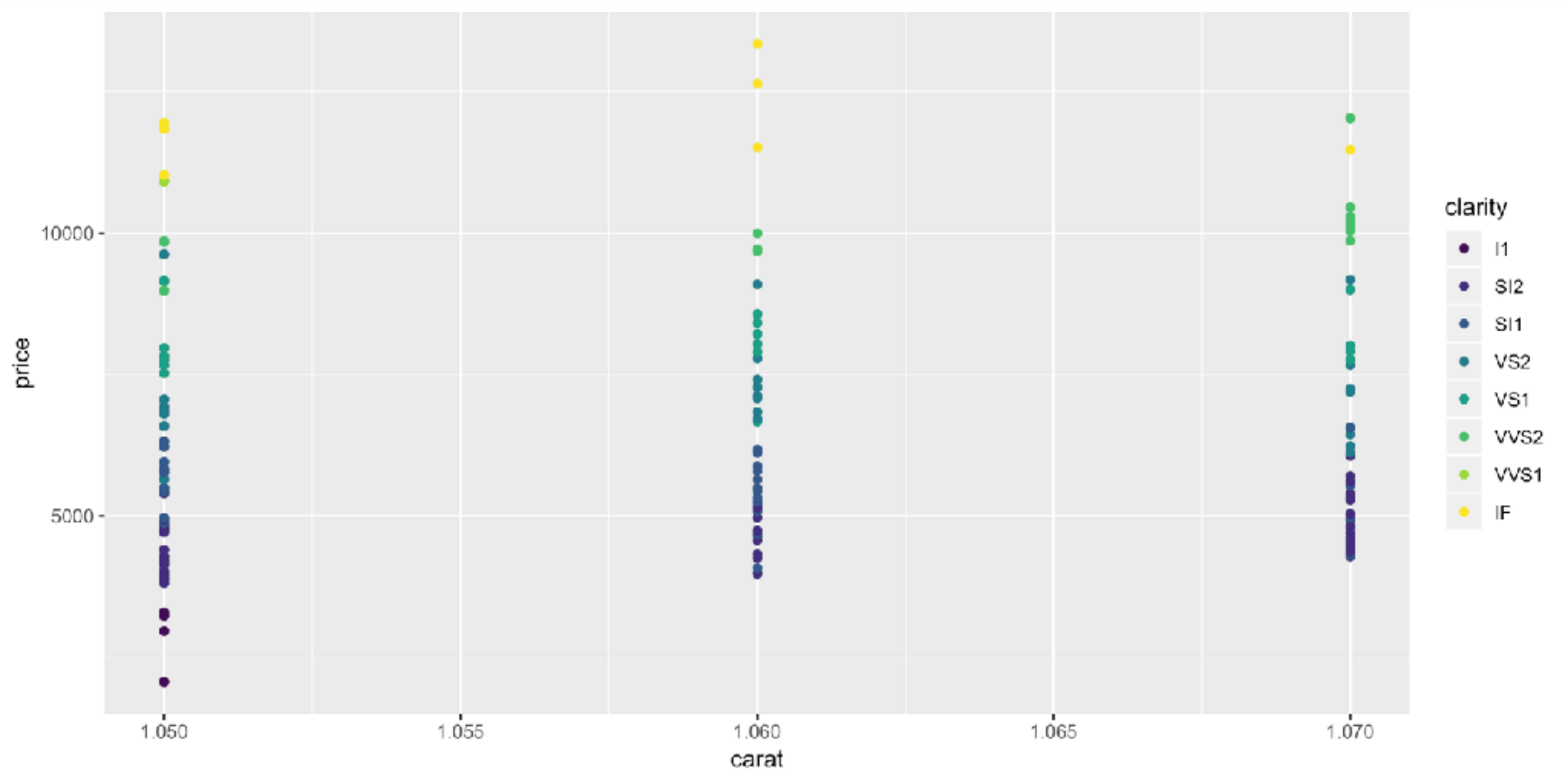
Bar plot of diamond prices in a small range of carats



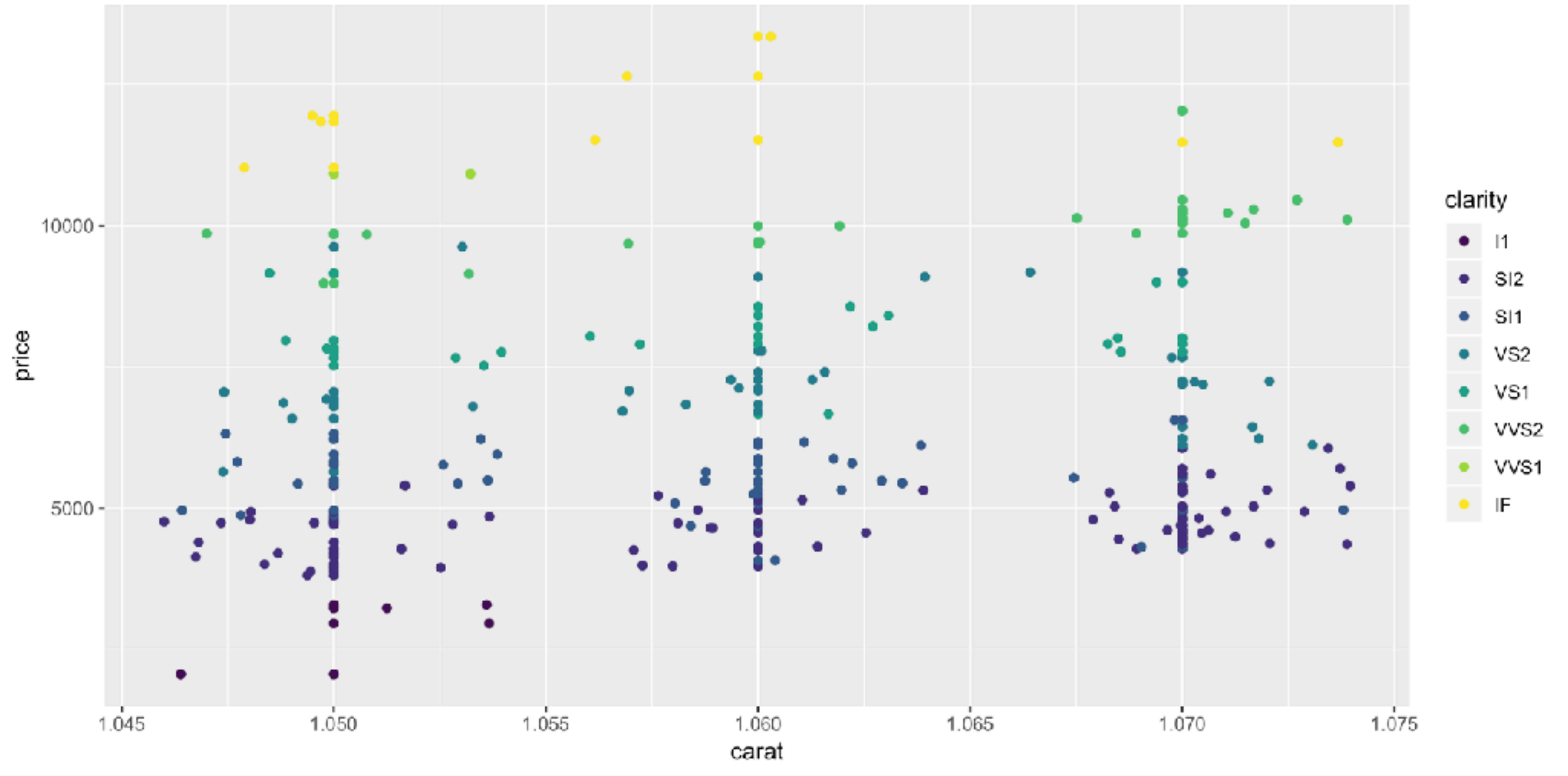
Scatter plot of diamond prices vs carat is same range



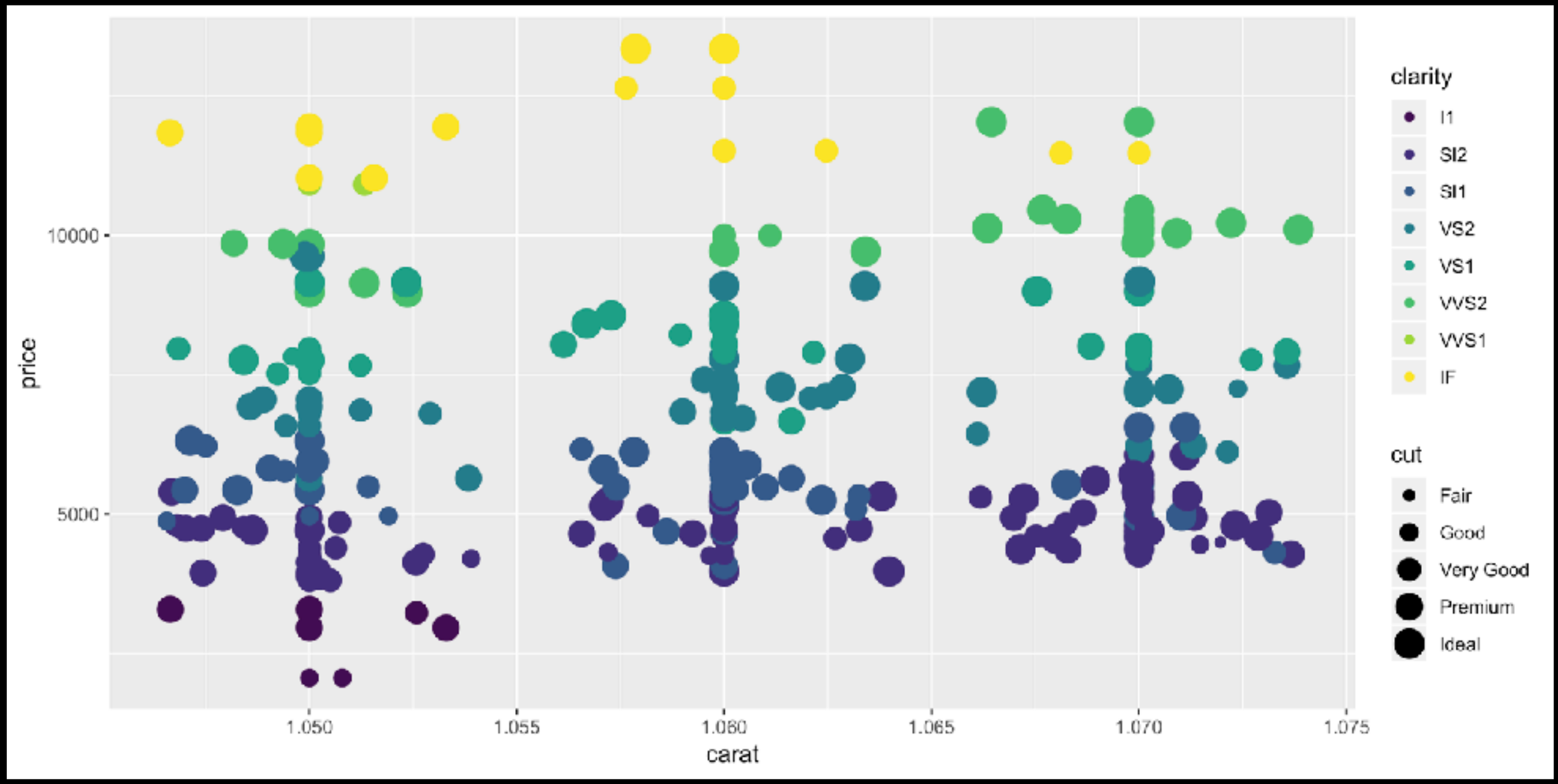
Scatter plot of same data now coloured according to clairity



Points “jittered” to see pattern



Size of points varied according to cut



Can also be redundant in the channels used to accentuate relationships.

e.g. using colour and size for clarity

Limitations

Cannot “mix and match” channels and marks entirely arbitrarily

Example - cannot adjust length of lines as already `taken’

Can’t adjust bar plot height in diamond data

Adjust width?

Channel types

Two kinds of sensory modalities

Identity channels - *what* something is; *where* it is

What-where (metathetic)

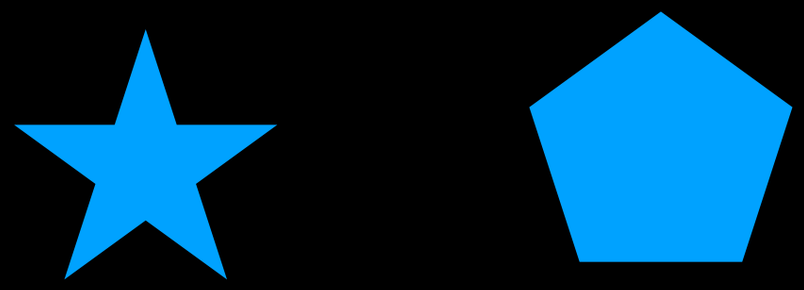
Magnitude channels - *how much* there is

How much (prothetic)

Examples

Identity channels

Shape



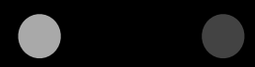
Colour hue

Can tell where the mark is on e.g. a grid, geometry or map

Magnitude channels

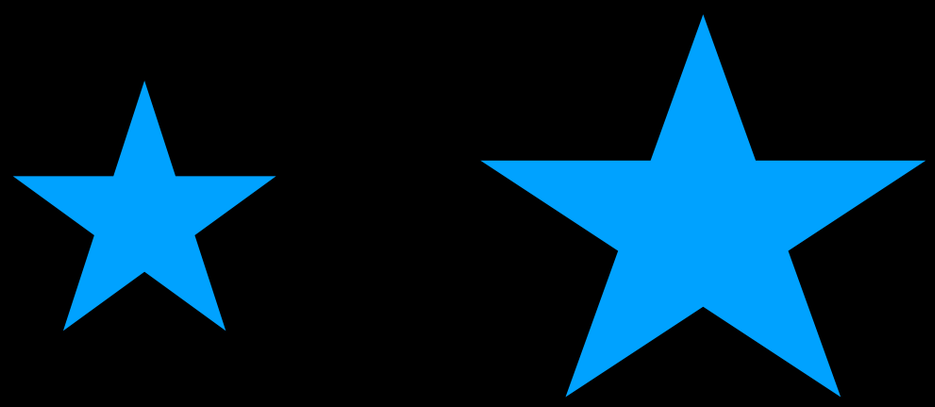
Two lines - which is longer?

Brightness



Tilt

Area



Channels are not all equally useful.

Want channels to be

Expressive

Effective

Expressiveness principle

the visual encoding should express all of, and only, the information in the dataset attributes

Ordered data should be seen in a way we perceive as ordered

Unordered data should *not* be seen in a way we perceive as ordered

Example - Not Expressive

Representing a set of heights with a set of symbols

< 1.0 m - Use a Star

> 1.0 m and < 2.0 m - Use a plus

Effectiveness principle

The most important attributes should be encoded with the most *noticeable* channels.

Decreasingly important attributes are matched with less *noticeable* channels.

Noticeable ?

Our brains do not notice different channels to the same level

Channels: Expressiveness Types and Effectiveness Ranks

Magnitude Channels: Ordered Attributes

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

Same

Same

Most ▲ Effectiveness ▼ Least

Identity Channels: Categorical Attributes

- Spatial region
- Color hue
- Motion
- Shape

Aspects of effectiveness

Accuracy

How close is human perceptual judgment to some objective measure?

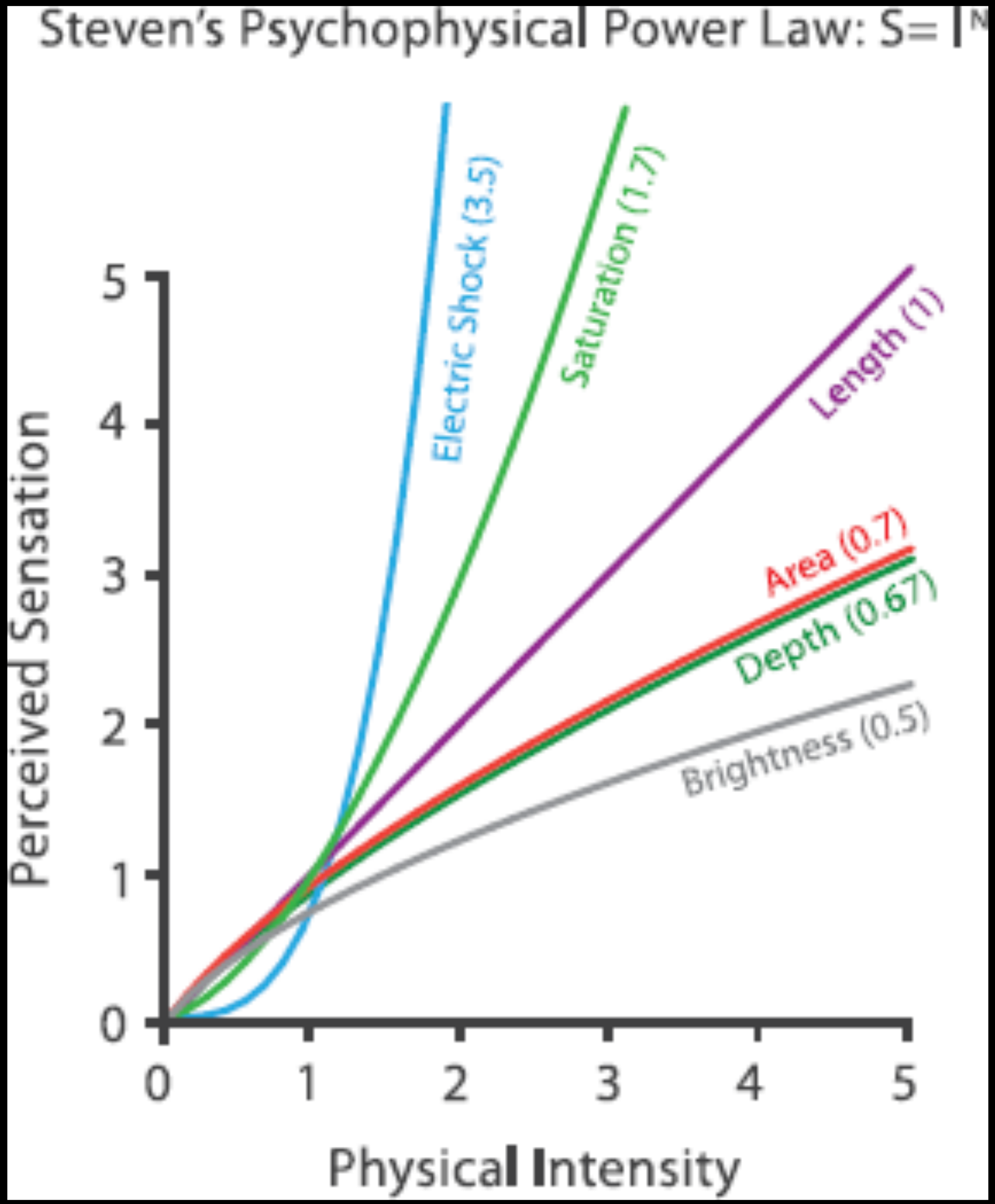
Relatively simple Mathematical model

Steven’s Psychophysical Power Law

$$S \propto I^n$$

S - Perceived Sensation

I - Physical Intensity


Steven's Psychophysical Power Law: S= I^N
Perceived Sensation
Physical Intensity
0
1
2
3
4
5
Electric Shock (3.5)
Saturation (1.7)
Length (1)
Area (0.7)
Depth (0.67)
Brightness (0.5)

Nobody’s ever figured out to use Electric Shock in Visualisation

Well almost nobody

Heer and Bostock, Crowd-sourced effort 2010

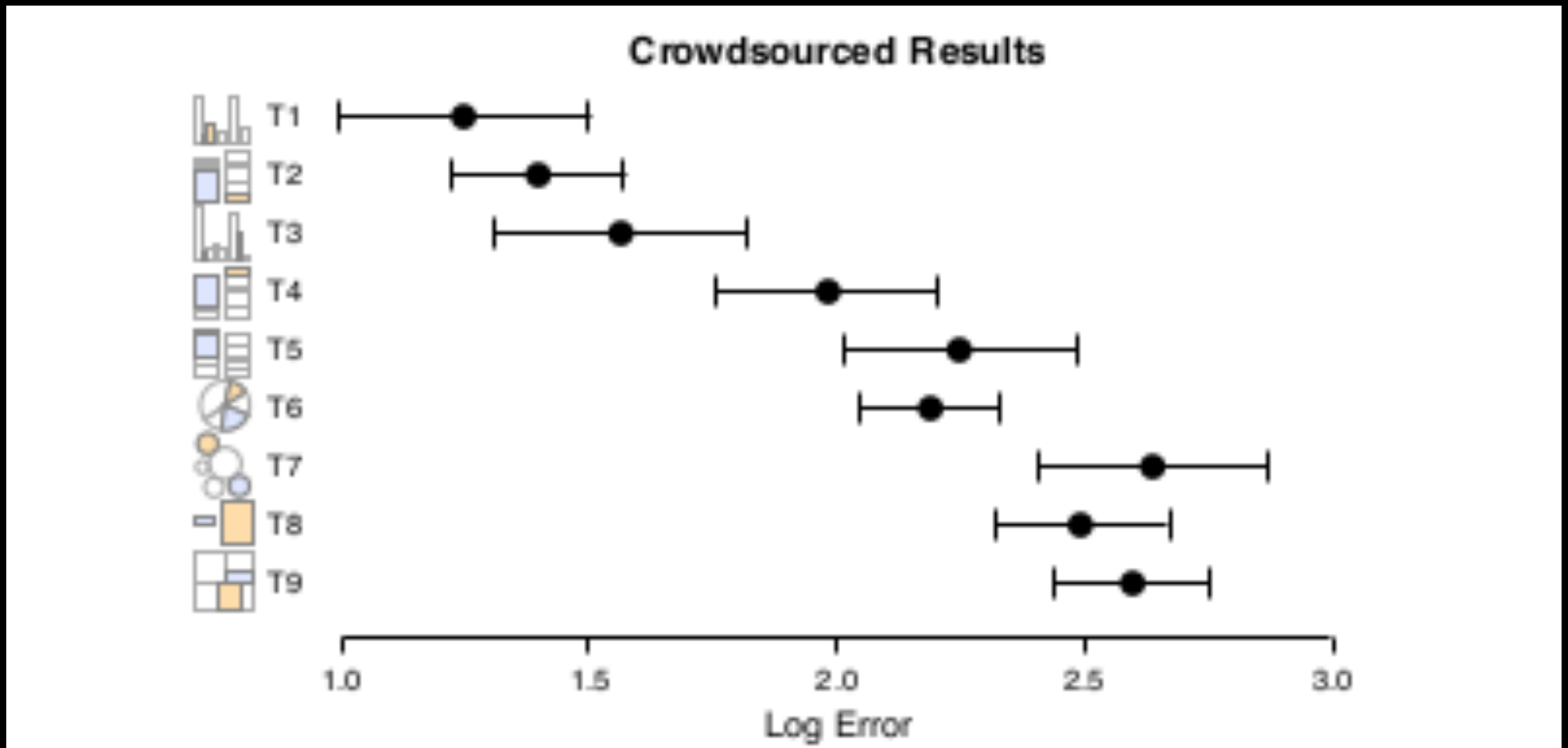


Figure 4: Proportional judgment results (Exp. 1A & B). Top: Cleveland & McGill's [7] lab study. Bottom: MTurk studies. Error bars indicate 95% confidence intervals.

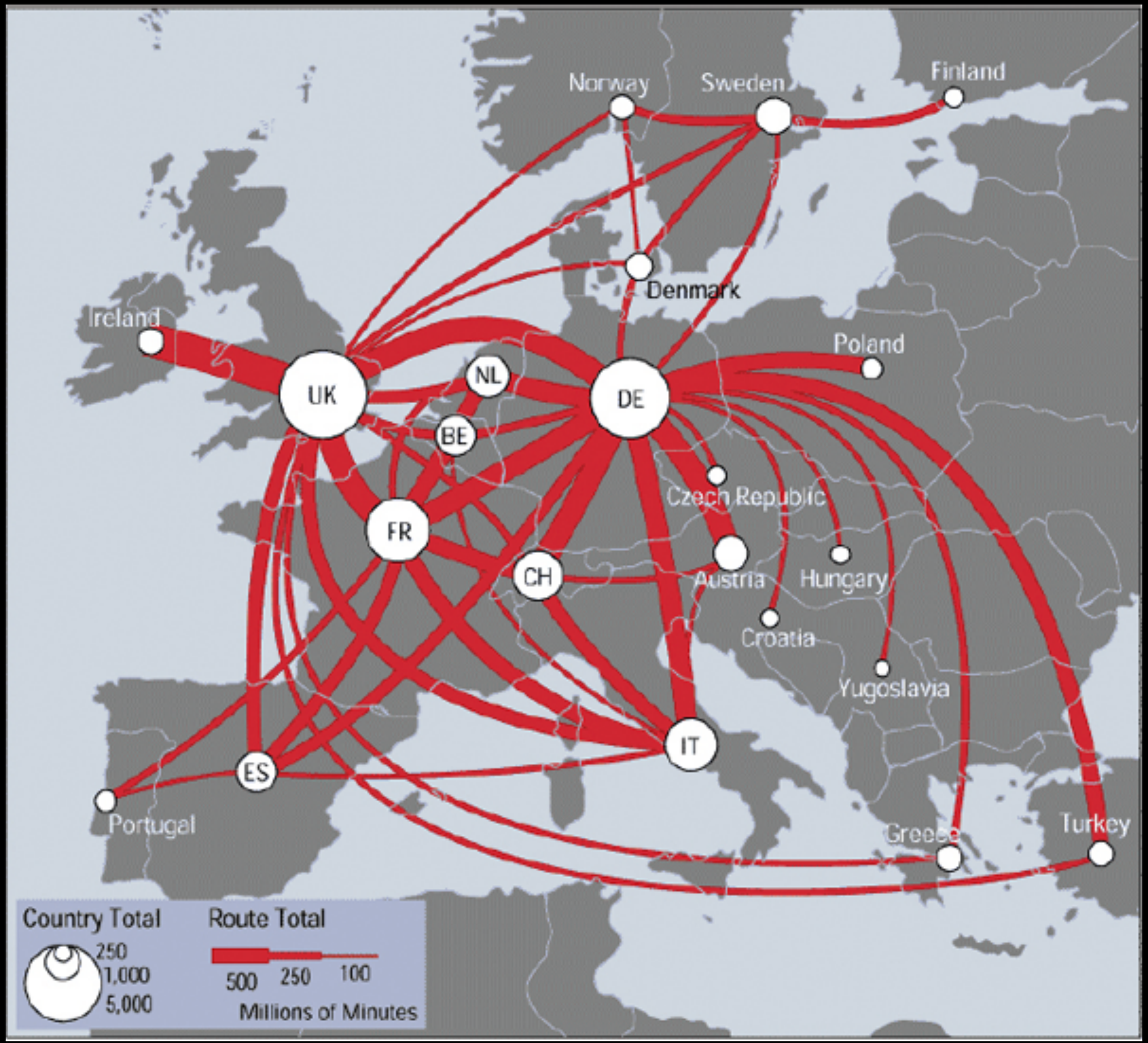
Discriminability

How perceptible are differences between items to a human?

Example width of a line

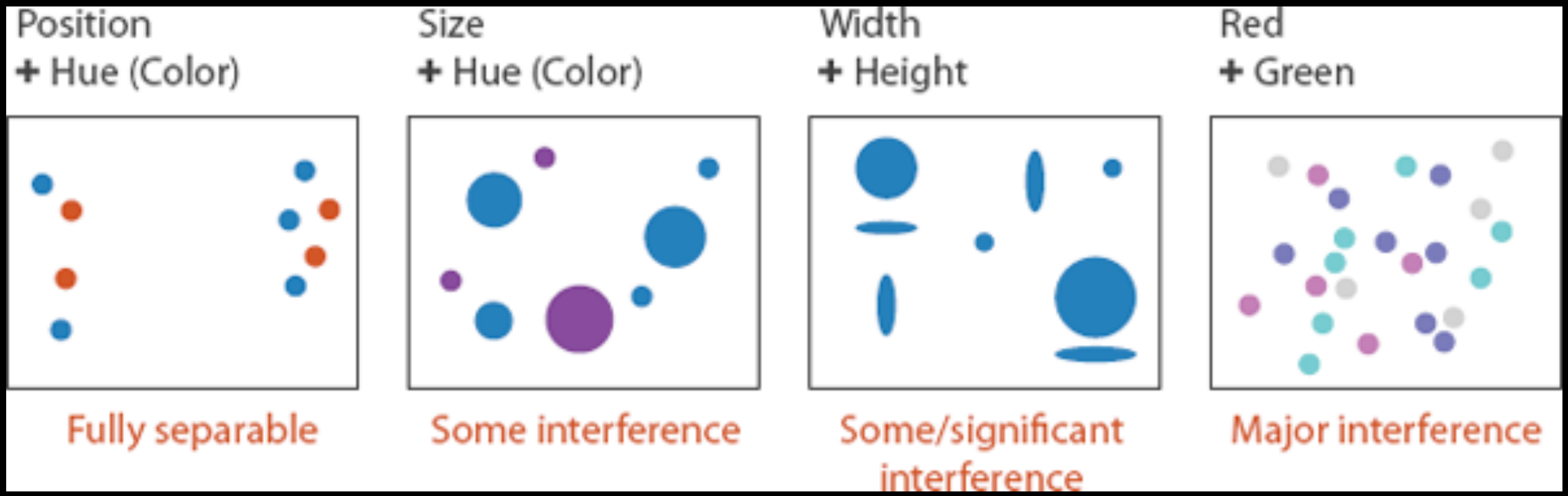
Good for three to four ordered values

Not for hundreds!


Norway
Sweden
Finland
Denmark
Ireland
Poland
UK
NL
DE
BE
Czech Republic
FR
CH
Austria
Hungary
Croatia
Yugoslavia
IT
ES
Portugal
Greece
Turkey
Country Total
250
1,000
5,000
Route Total
500
250
100
Millions of Minutes

Separability

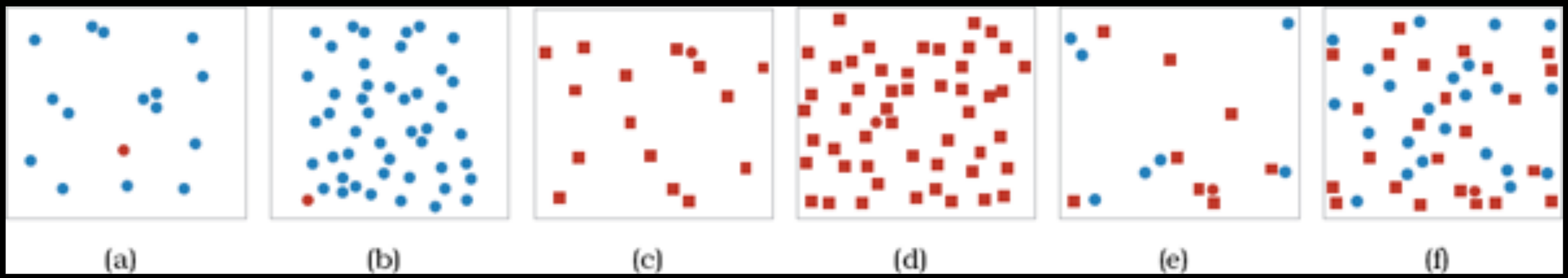
Channels can interact with each other



Popout

How well a distinct item stands out from many others immediately.

Colour much better than shape.



Weber’s Law (1834)

Human perception based on relative, not absolute, changes

Example (taken from https://bit.ly/2E0HgNa)

Suppose you are lifting weights and you are blindfolded

Initially holding 2 Kg

How much more needs to be added before you notice a change?

Suppose you only notice a change after extra 0.2 Kg added.

What if you are now holding 5 Kg - what is smallest noticeable change?

Find only notice a change after extra 0.5 Kg added.

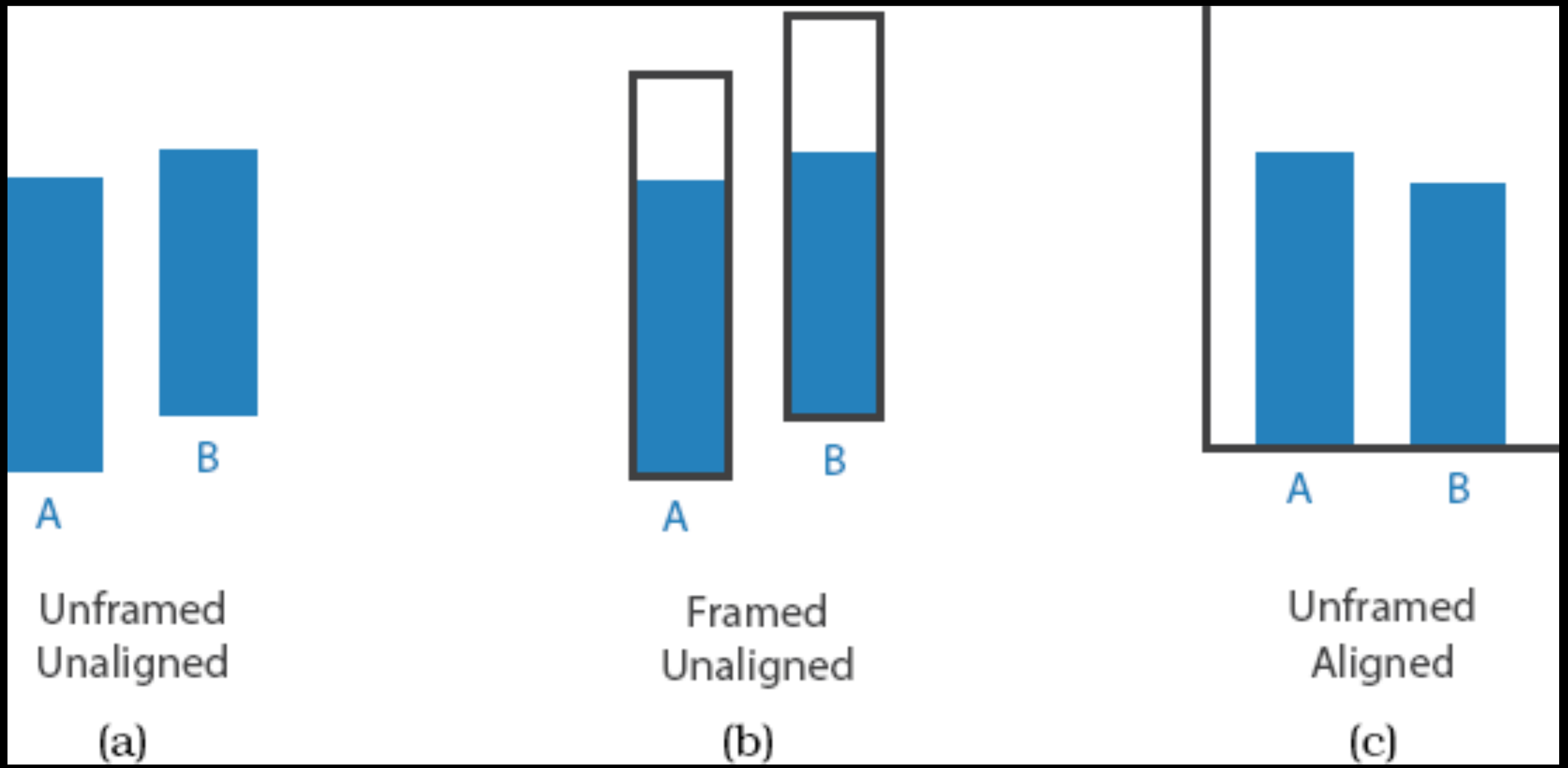
$$\frac{\text{Smallest Noticeable Change}}{\text{Intensity}} = \text{Constant}$$

$$\frac{\text{Smallest Noticeable Change}}{\text{Intensity}} = \text{Constant}$$

Weber's Law

Holds also for channels in visualisation

Upshot - need to think carefully about how to relative comparisons



Unframed
Unaligned

(a)

Framed
Unaligned

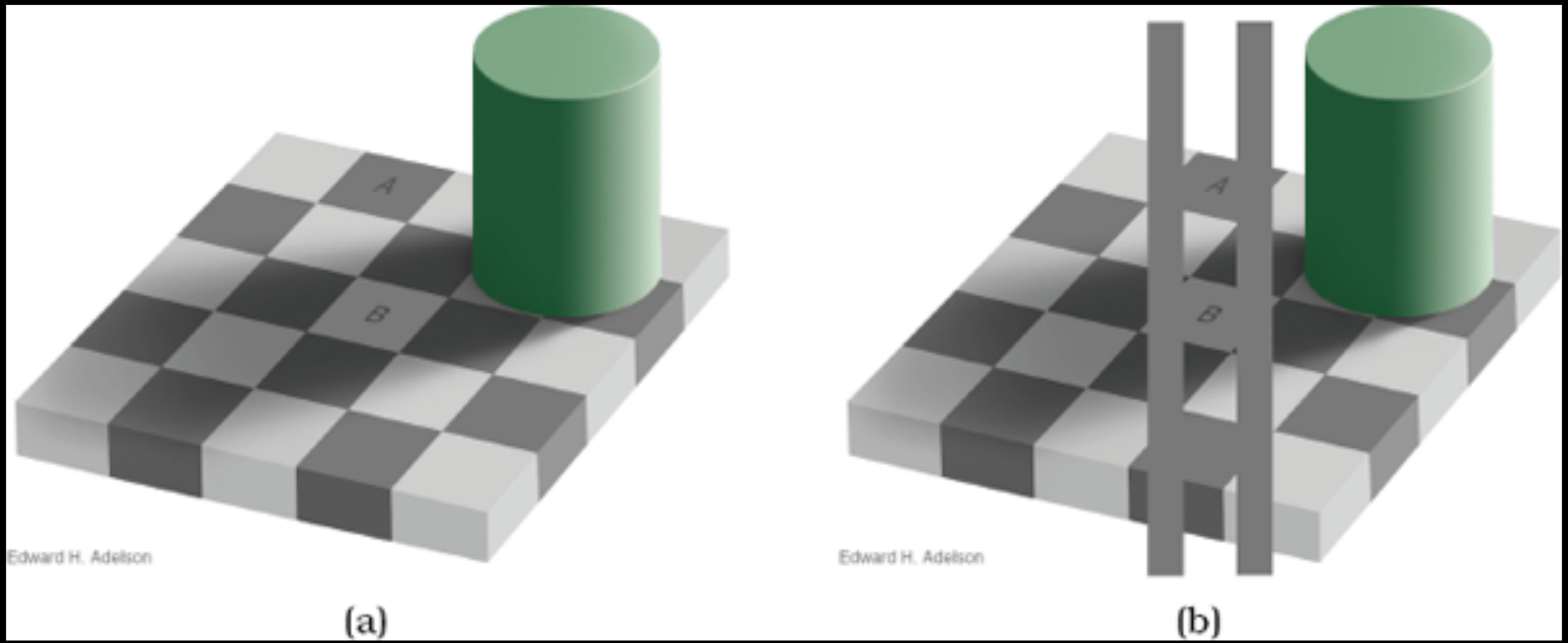
(b)

Unframed
Aligned

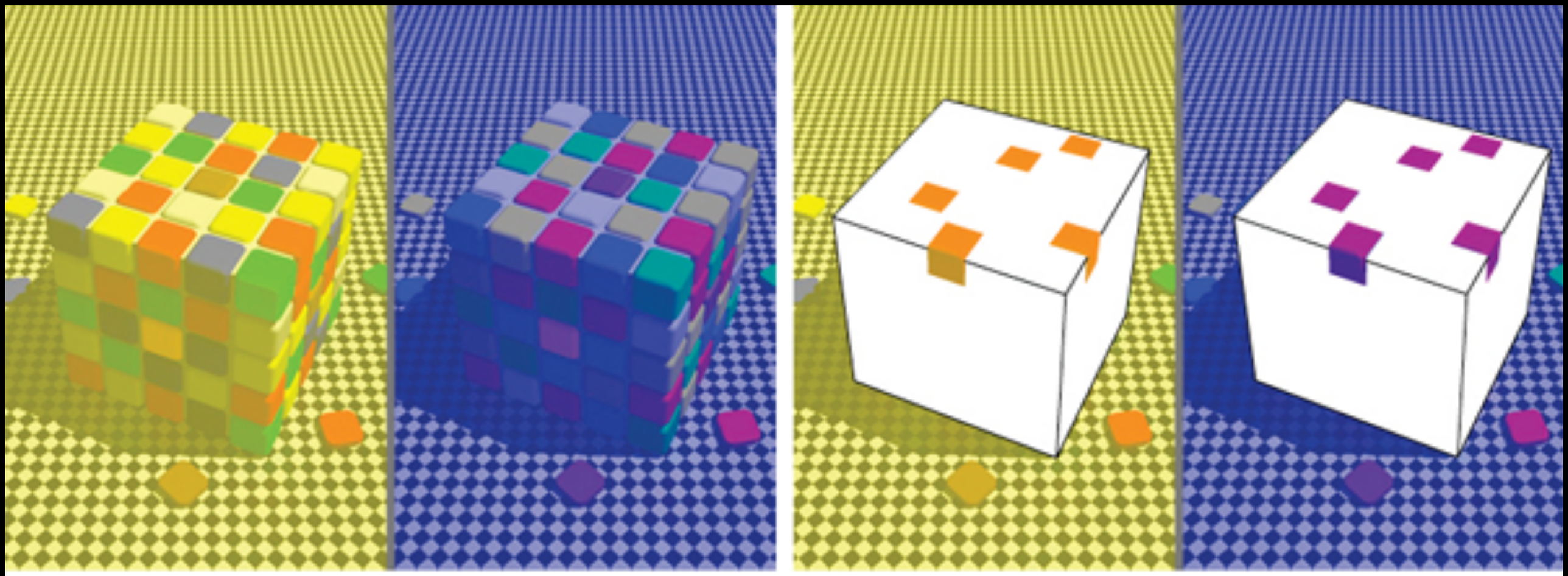
(c)

Which is longer A or B?

Situation even harder with colour and contrast….



(a) (b)



(a) (b)

Rules of Thumb

No unjustified 3d

No unjustified 3d

Eyes beat memory

No unjustified 3d

Eyes beat memory

Resolution ever Immersion

No unjustified 3d

Eyes beat memory

Resolution ever Immersion

Overview First, Zoom and Filter, Detail on Demand

No unjustified 3d

Eyes beat memory

Resolution ever Immersion

Overview First, Zoom and Filter, Detail on Demand

Responsiveness is Required

No unjustified 3d

Eyes beat memory

Resolution ever Immersion

Overview First, Zoom and Filter, Detail on Demand

Responsiveness is Required

Get in right in Black & White

No unjustified 3d

Eyes beat memory

Resolution ever Immersion

Overview First, Zoom and Filter, Detail on Demand

Responsiveness is Required

Get in right in Black & White

Function first, Form next

No unjustified 3d

Why?

Deal with 3d all our lives

Reasons - Depth perception is poor

Remember Steven’s Law - Depth perception less accurate than area perception

Occlusion hides information

Used in painting as a powerful way to infer depth

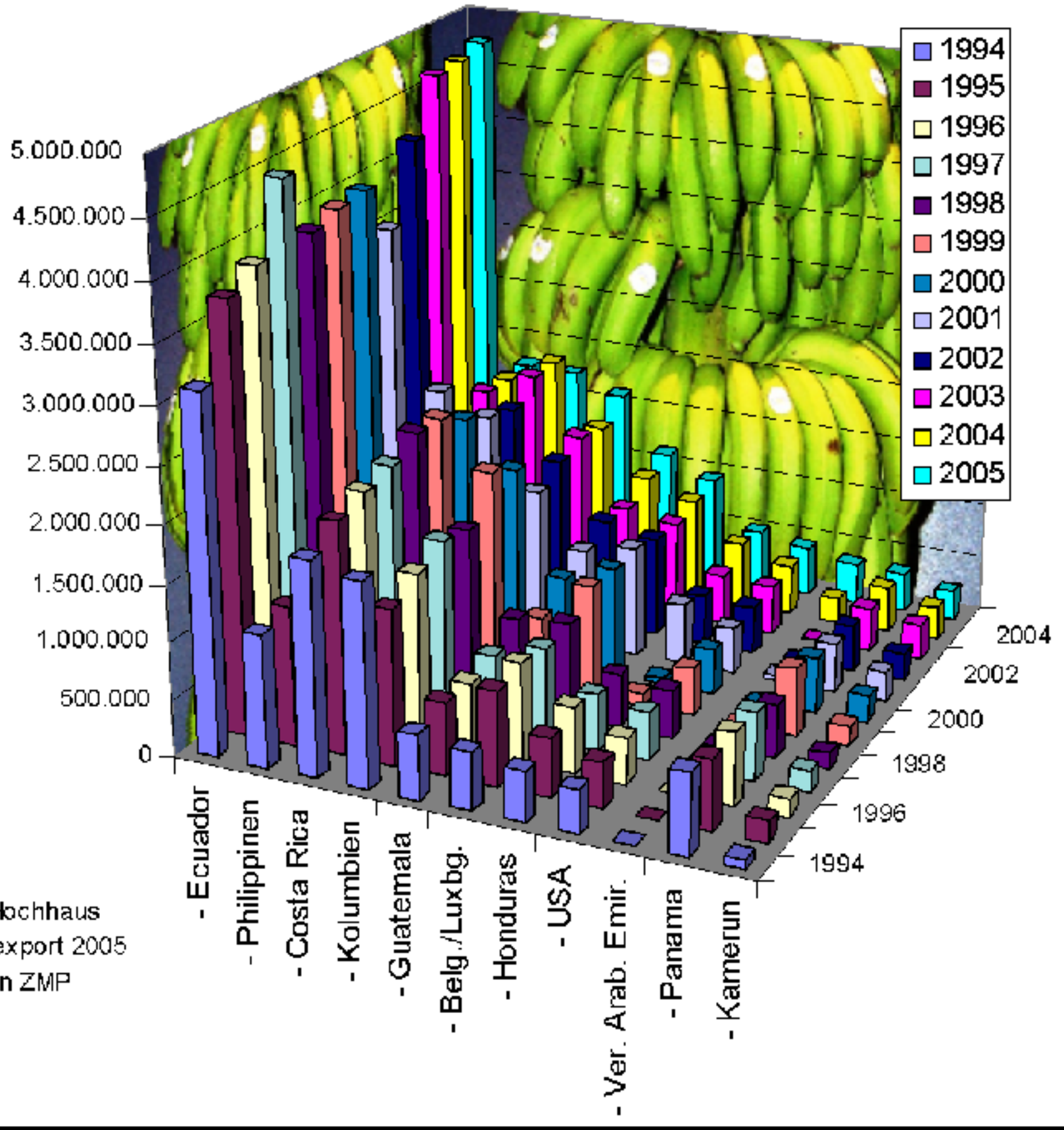
Motion parallex - changing occlusion as we change viewing position

Interactive 3d tools allow us to simulate motion parallex

Example Protein Structure

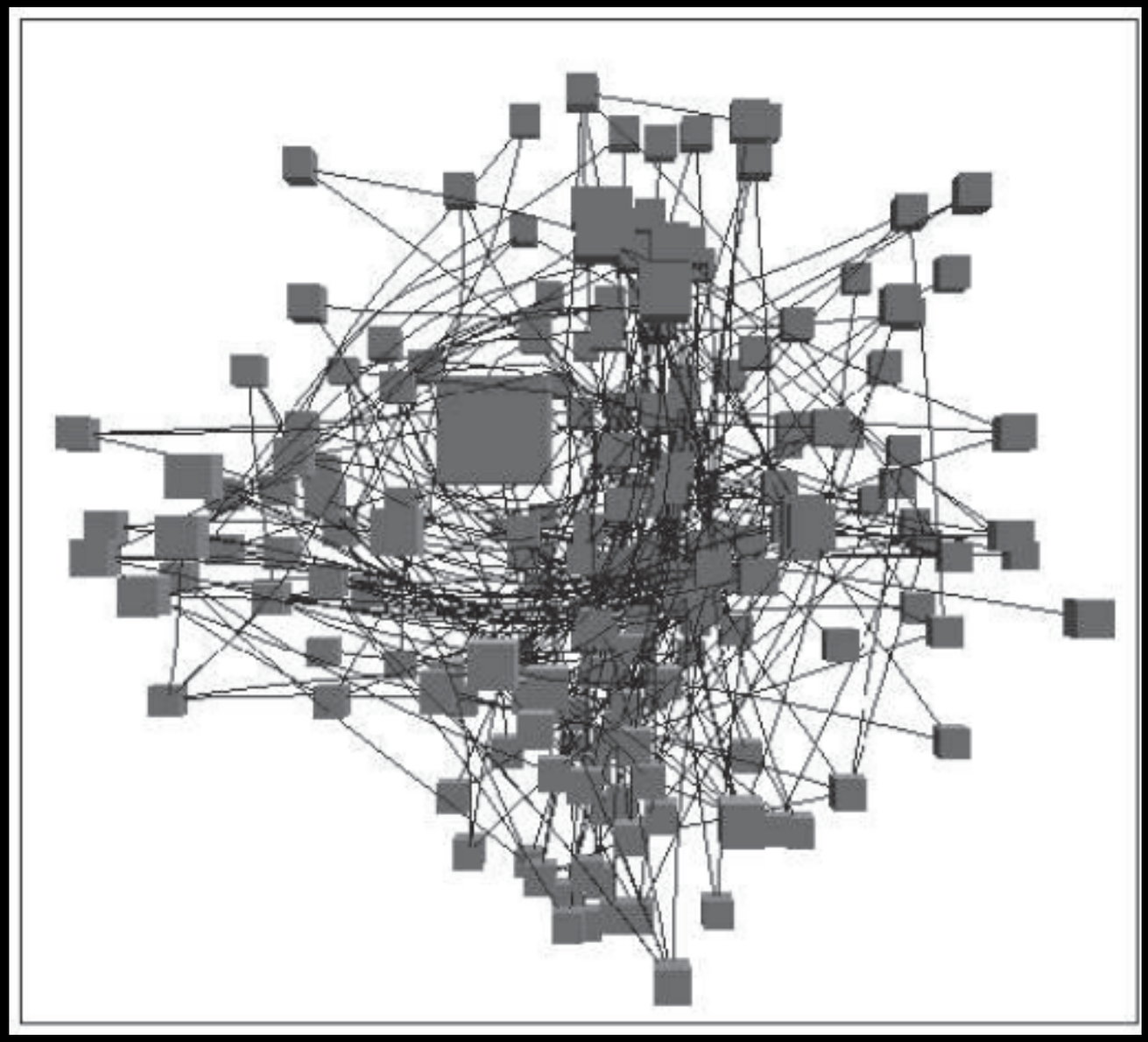
If you have static image - occlusion hides information!

Export von Bananen in Tonnen von 1994-2005



Dr. Hochhaus
Banexport 2005
Daten ZMP

Even if you have an interactive tool…



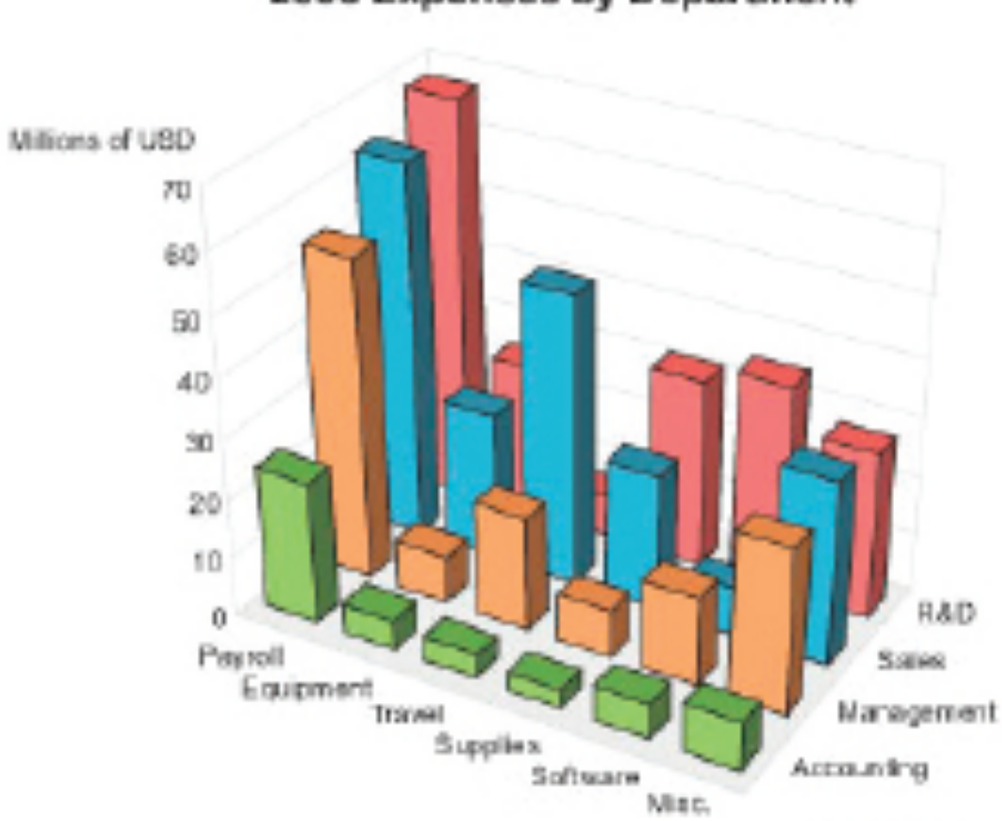
Many 3d structures are unfamiliar

Perspective distorts distances

Ancient Greeks adjusted widths of temples so that they “look correct” from a distance

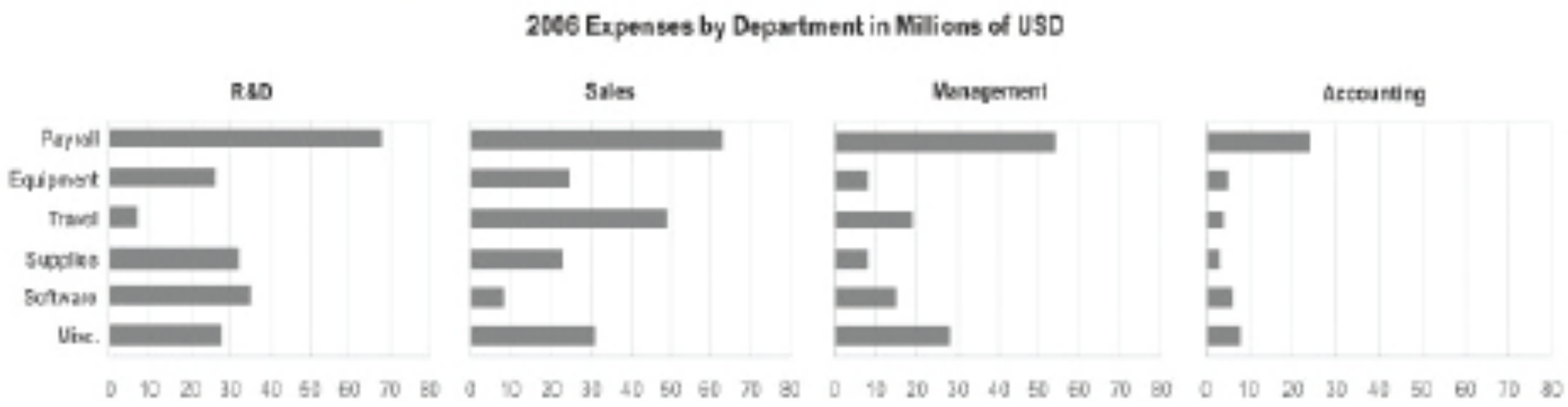
Graph Design I.Q. Test

Question 7: Which graph makes it easier to determine R&D's travel expense?



3-D Bar Graph (left)

2-D Bar Graphs (below)



Other depth cues

- Shadows
- Comparison with object of known size (e.g. car)

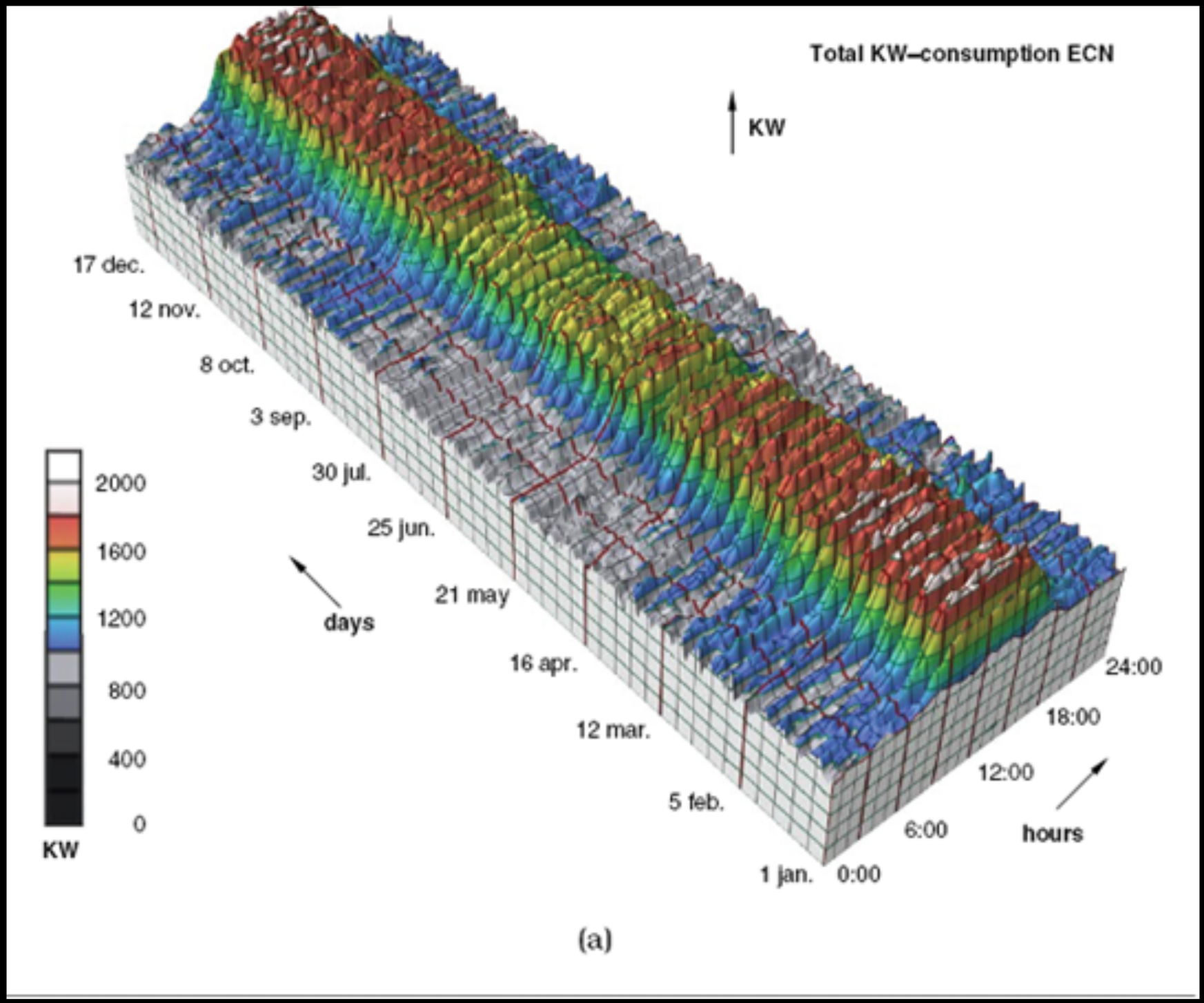
Well maybe not for everybody

Stronger than stereo-vision depth

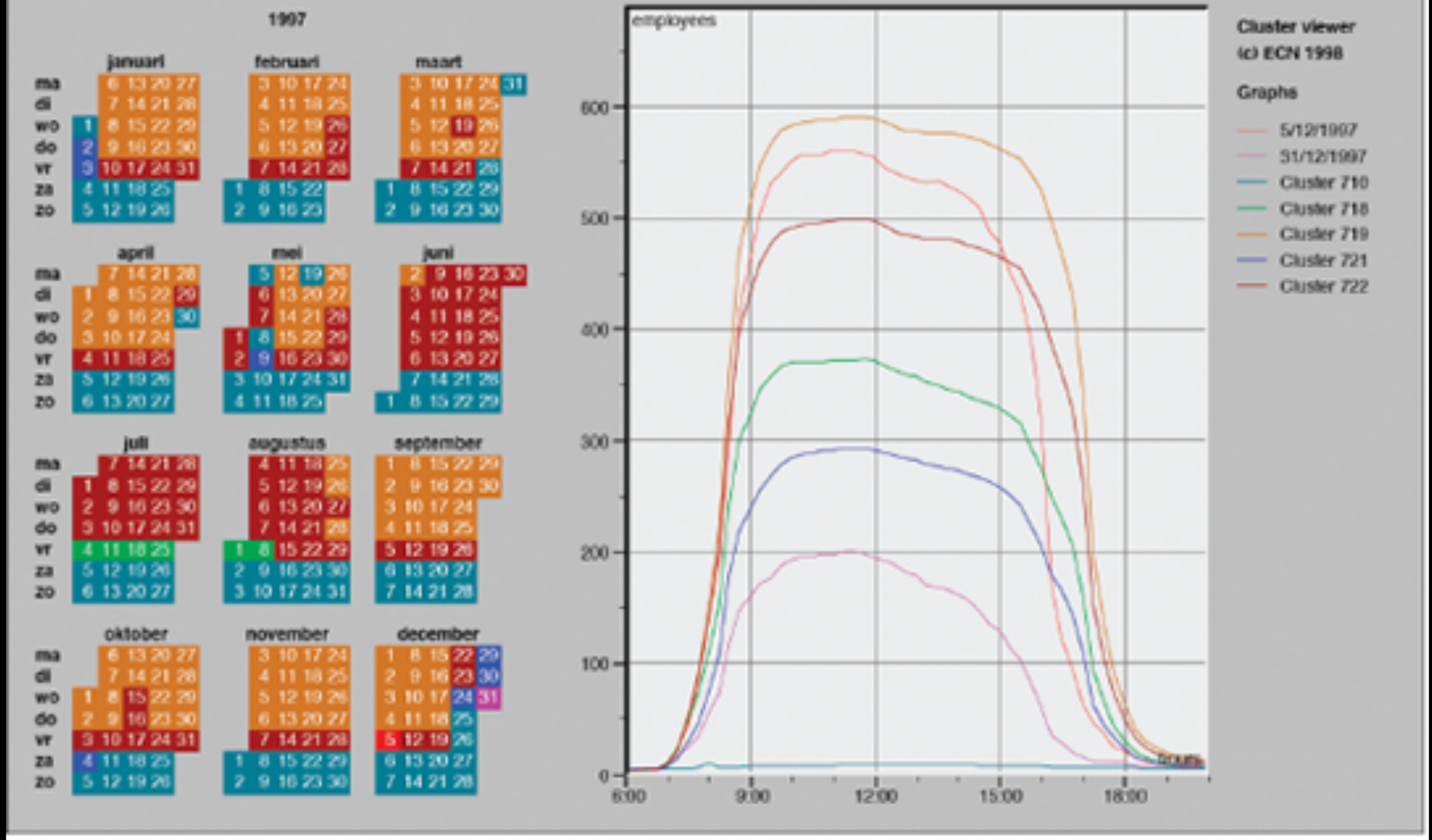
Adds clutter to image !

Shadows etc. affect colour

What do you do with 3D looking data ?



(a)

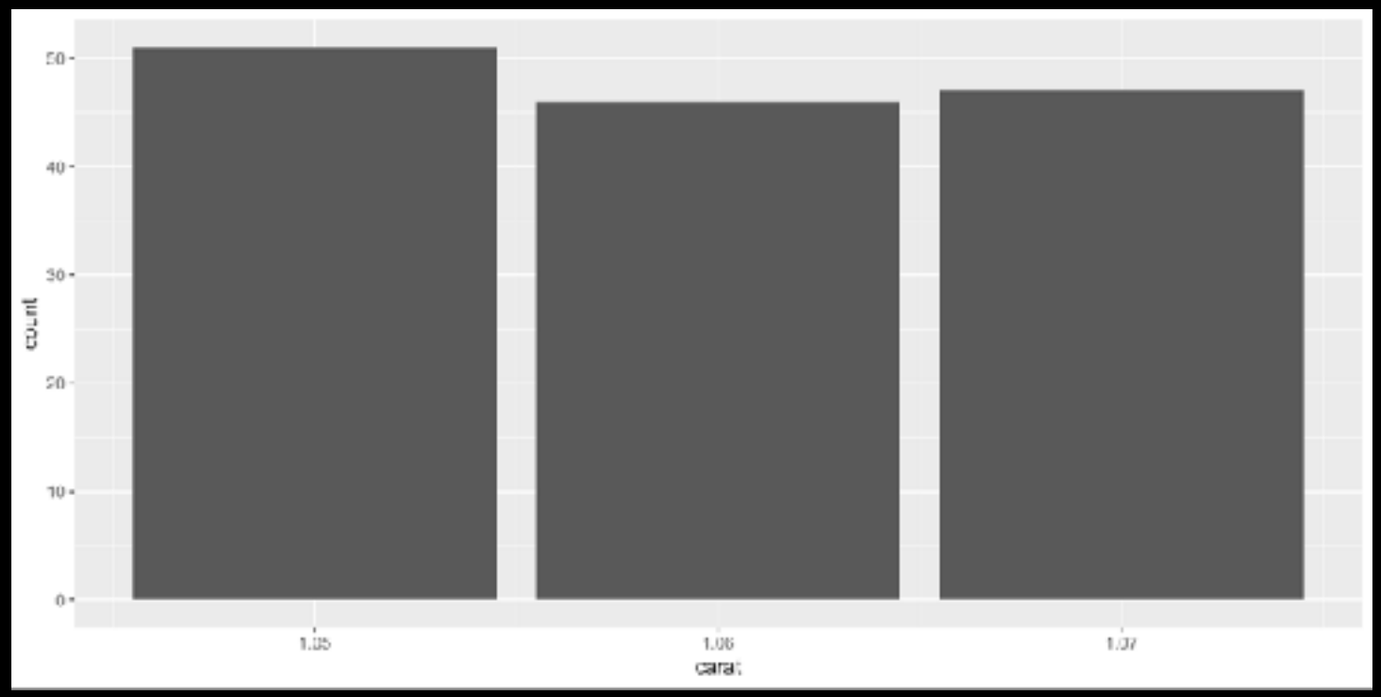

1997
januari
februari
maart
april
mei
juni
juli
augustus
september
oktober
november
december
ma
di
wo
do
vr
za
zo
employees
600
500
400
300
200
100
0
6:00
9:00
12:00
15:00
18:00
Cluster viewer
(c) ECN 1998
Graphs
5/12/1997
31/12/1997
Cluster 710
Cluster 718
Cluster 719
Cluster 721
Cluster 722

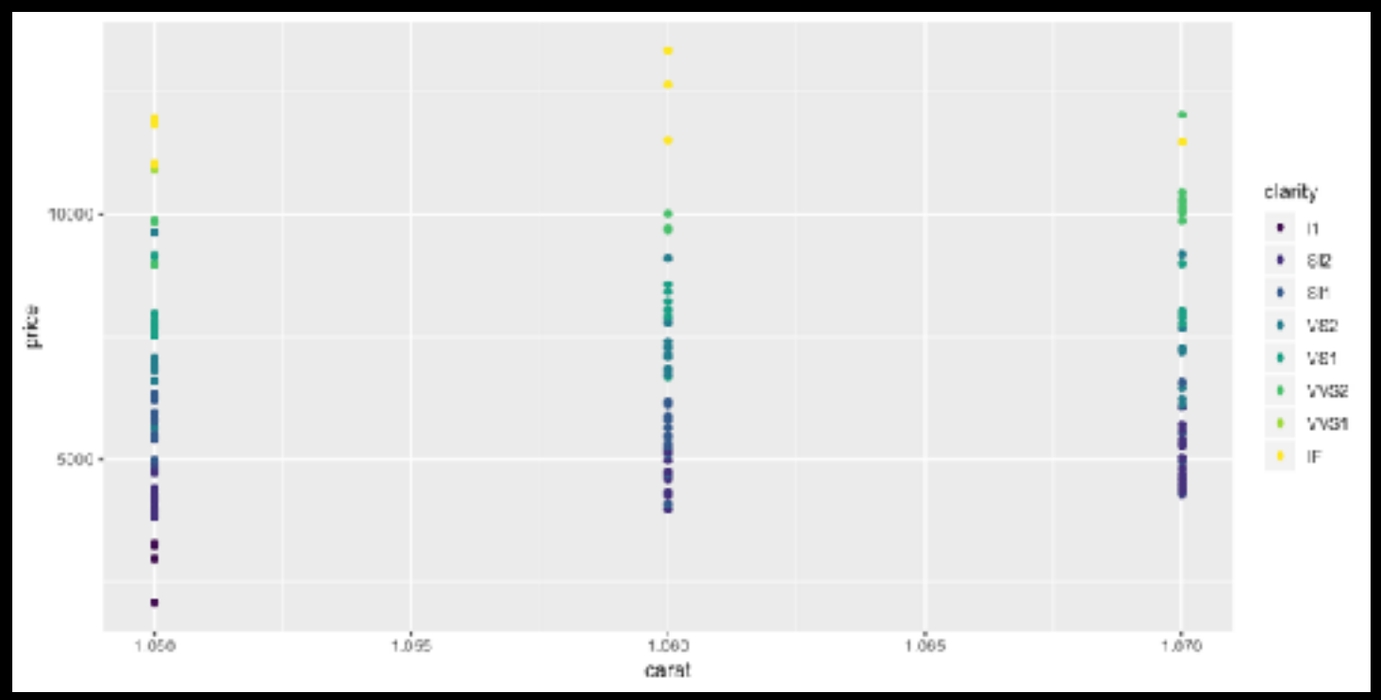
(b)

Eyes beat memory

Switching between different views that are visible simultaneously
has a much lower cognitive load
than using our memories.

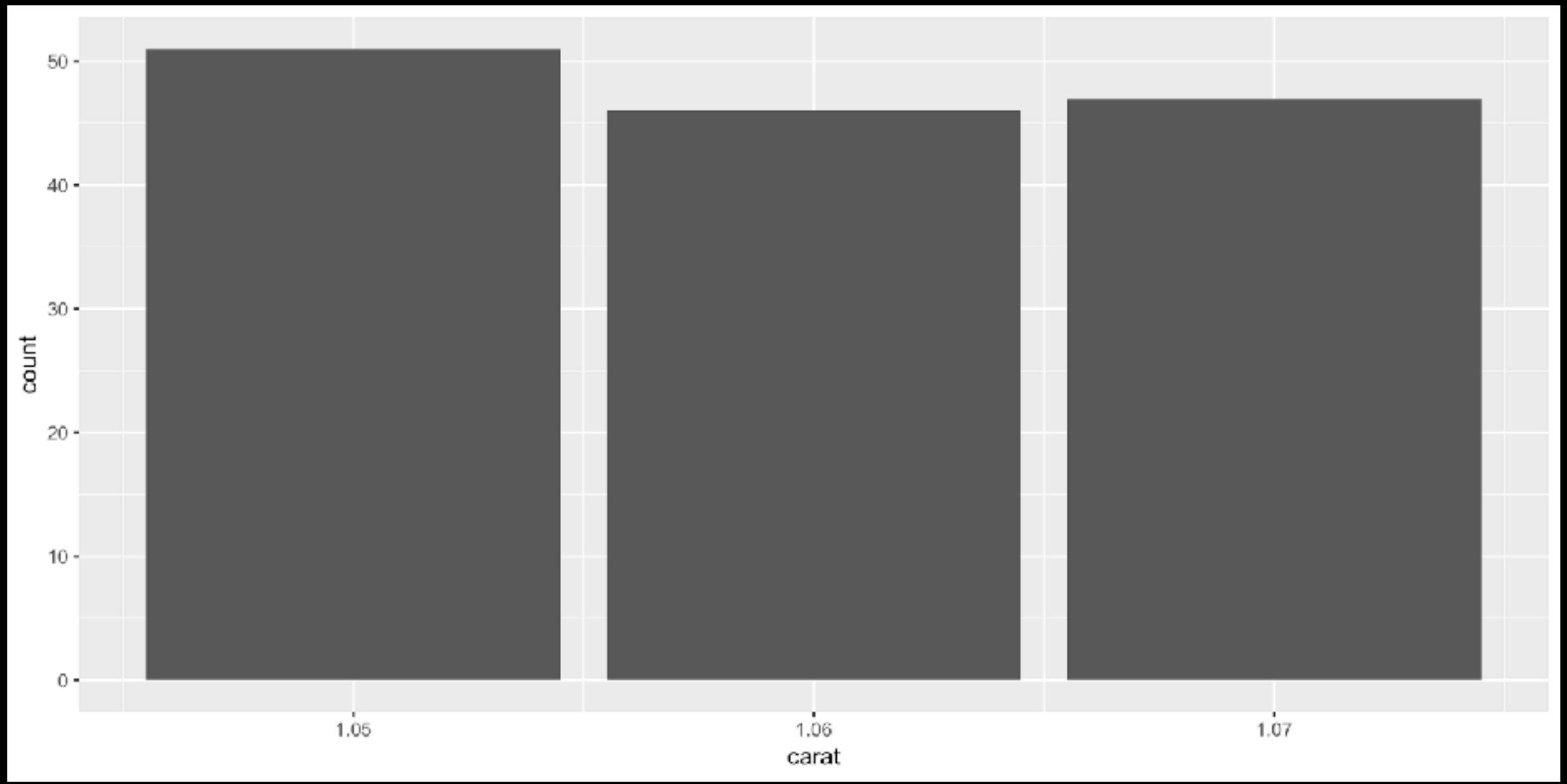
In other words, this ….



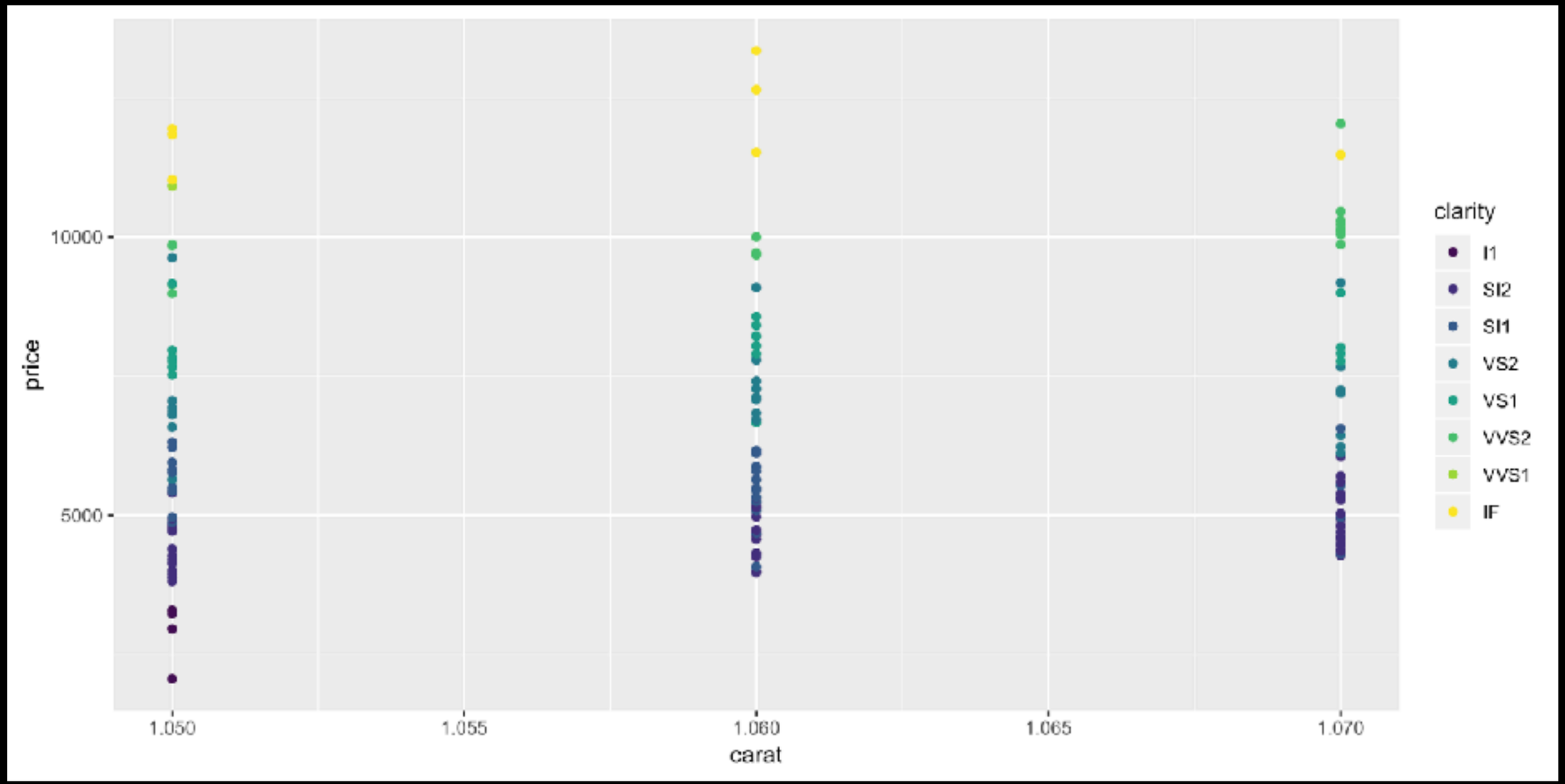


Is better than this…

Is better than this…



Is better than this…



Two categories of memory

Short term - lasts several seconds - working memory

Long term - can last a lifetime

Short term memory - limited

Cognitive load when reach limits

Fail to absorb information

Facet plots are useful

Animation

Important tool but caution…

Good for showing transitions

Multiple animated frames can lose detail in comparison with frames

Year: 1952

Africa | Americas | Asia

Europe | Oceania

life expectancy

80
60
40

1e+03 1e+04 1e+05

GDP per capita

Year: 1952

Africa | Americas | Asia

Europe | Oceania

life expectancy

80
60
40

1e+03 1e+04 1e+05

GDP per capita

Change blindness

Get it right in Black and White

Noted effectiveness of colour

Check most crucial aspects of visualisation there in Black & White

Don’t print in colour as a check

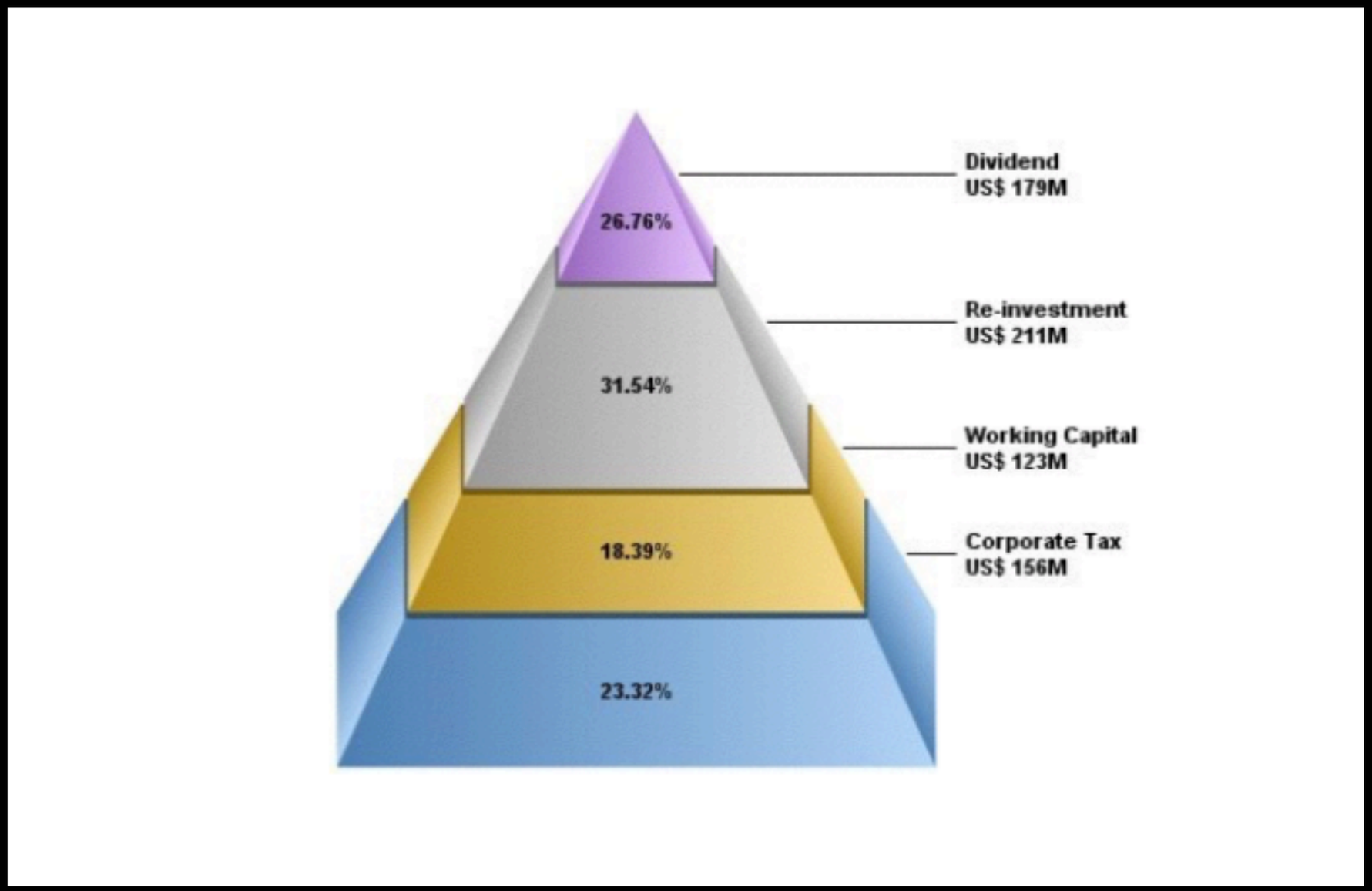
Function First, Form Next

Concentrate first on telling the story, then on making it look good.

Bad Graphics!

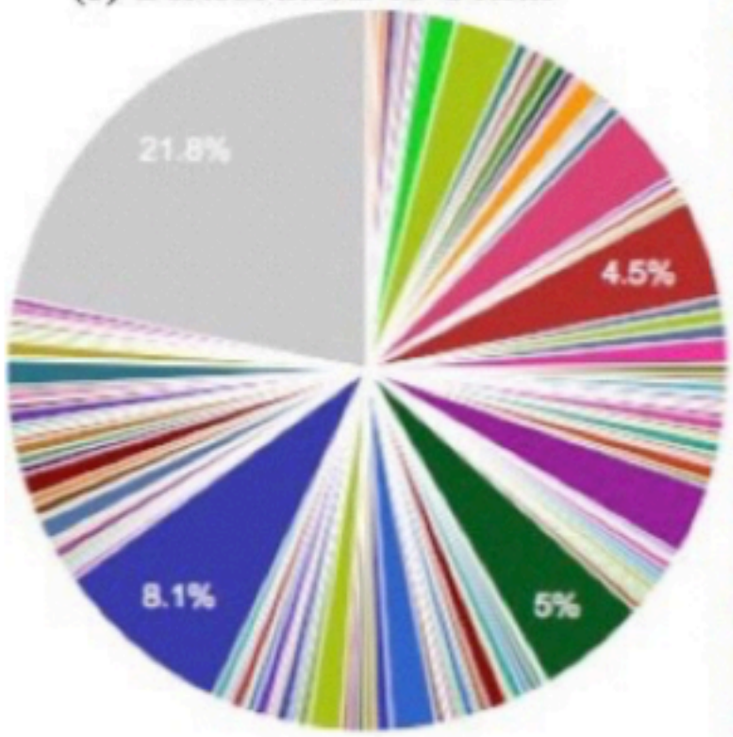
What’s wrong with these images?

How would you plot it instead?

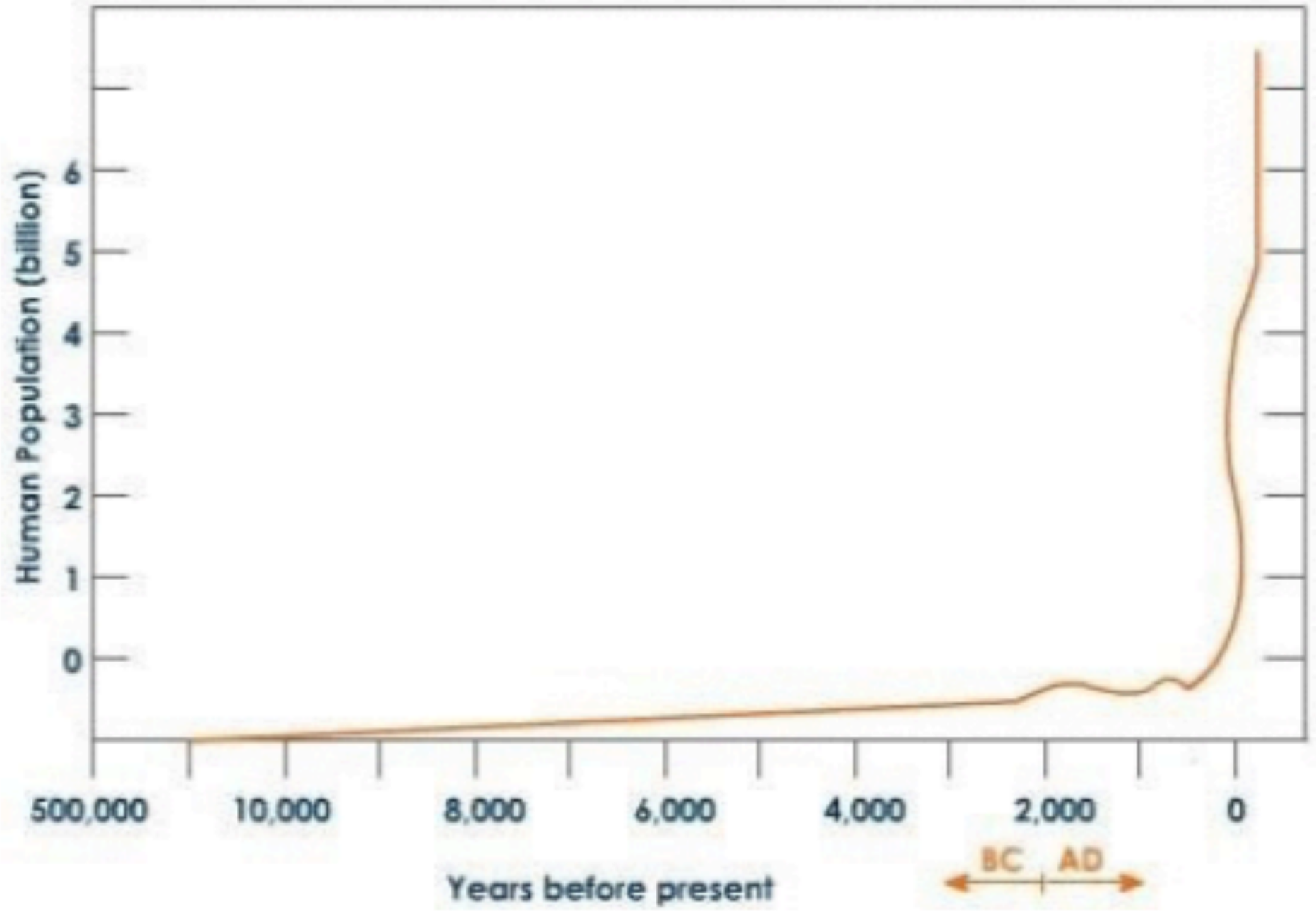

26.76%
31.54%
18.39%
23.32%
Dividend
US$ 179M
Re-investment
US$ 211M
Working Capital
US$ 123M
Corporate Tax
US$ 156M

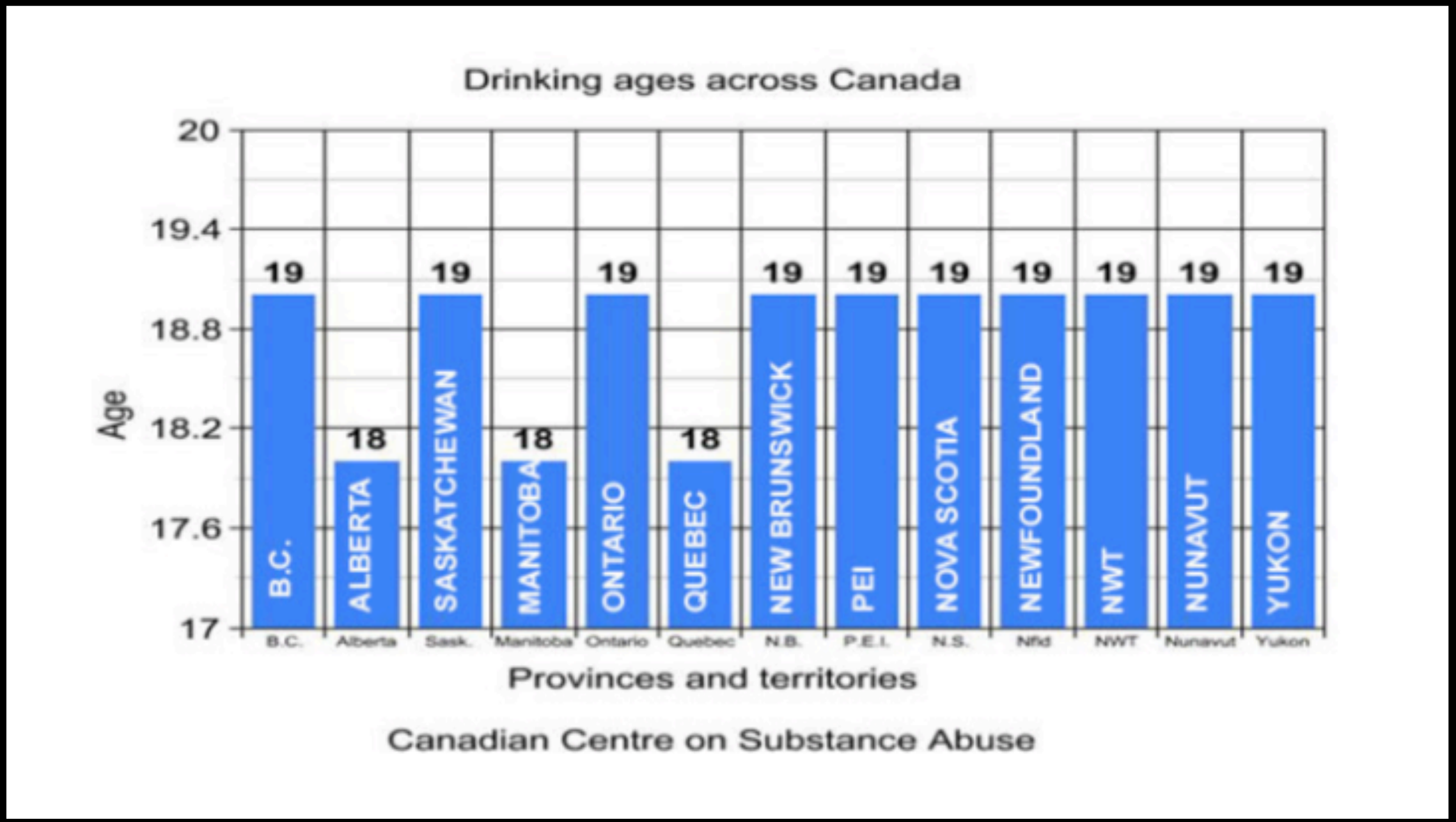

KNUCKLEBALL VELOCITY
R.A. DICKEY
77.3 MPH
75.3 MPH
2012
2013
BlackBerry
Keep Moving
BLUE JAYS

(f) Distribution of Genus



- Acaryochloris
- Achromobacter
- Acidiphilium
- Acidobacterium
- Acidovorax
- Acinetobacter
- Albidiferax
- Algoriphagus
- Alkalilimnicola
- Allochromatium
- Ammonifex
- Anabaena
- Anaerolinea
- Anaeromyxobac
- Aromatoleum
- Arthrospira
- Azoarcus
- Bacillus
- Bacteroides
- Bordetella
- Bradyrhizobium
- Burkholderia
- Caldanaerobacter
- Caldicellulosiruptor
- Candidatus Accumulib
- Candidatus Desulforuc
- Candidatus Koribacter
- Candidatus Solibacter
- Carboxydothermu
- Caulobacter
- Chitinophaga
- Chlorobaculum
- Chlorobium
- Chloroflexus
- Chthoniobacter
- Clostridium
- Cupriavidus
- Cyanothece
- Cytophaga
- Dechloromonas
- Dehalococcoides
- Deinococcus
- Desulfarculus
- Desulfatibacillun
- Desulfitobacteriu
- Desulfobacteriur
- Desulfococcus
- Desulfotomaculu
- Desulfovibrio
- Desulfuromonas
- Dictyoglomus
- Dyadobacter
- Flavobacterium
- Frankia
- Gemmata
- Geobacillus
- Geobacter
- Gloeobacter
- Haliangium
- Heliobacterium
- Herpetosiphon
- Isosphaera
- Ktedonobacter
- Leadbetterella
- Leptospira
- Leptothrix
- Lyngbya
- Magnetospirillum
- Methylacidiphilum
- Methylibium
- Methylobacterium
- Methylococcus
- Microcoleus
- Microcystis
- Microscilla
- Moorella
- Mycobacterium
- Myxococcus
- Nitrosococcus
- Nitrosomonas
- Nodularia
- Nostoc
- Opitutus
- Oscillatoria
- Oscillochloris
- Paenibacillus
- Pedobacter
- Pedosphaera
- Pelobacter
- Pelotomaculum
- Pirellula
- Planctomyces
- Pseudomonas
- Ralstonia
- Rhizobium
- Rhodopirellula
- Rhodopseudomc
- Rhodospirillum
- Rhodothermus
- Roseiflexus
- Roseomonas
- Rubrobacter
- Shewanella
- Siderovudans
- Sphingomonas
- Spirochaeta
- Spirosoma
- Stigmatella
- Streptomyces
- Symbiobacterium
- Synechococcus
- Synechocystis
- Syntrophobacter
- Syntrophus
- Thermaerobacter
- Thermincola
- Thermodesulfovibrio
- Thermomicrobium
- Thermosinus
- Thermosynechococcus
- Thermotoga
- Thermus
- Thioalkalivibrio
- Thiobacillus
- Trichodesmium
- Truepera
- Variovorax
- Xanthomor
- unclassifiec from Bacte
- unclassifiec from Deltaprotec
- Other


Human Population (billion)
6
5
4
3
2
1
0
500,000
10,000
8,000
6,000
4,000
2,000
0
BC
AD
Years before present


Drinking ages across Canada
20
19.4
18.8
18.2
17.6
17
Age
19
18
19
18
19
18
19
19
19
19
19
19
19
19
B.C.
ALBERTA
SASKATCHEWAN
MANITOBA
ONTARIO
QUEBEC
NEW BRUNSWICK
PEI
NOVA SCOTIA
NEWFOUNDLAND
NWT
NUNAVUT
YUKON
B.C.
Alberta
Sask.
Manitoba
Ontario
Quebec
N.B.
P.E.I.
N.S.
Nfld
NWT
Nunavut
Yukon
Provinces and territories
Canadian Centre on Substance Abuse