UNIVERSITY OF BIRMINGHAM

Visualisation

Week 1

Session 1

UNIVERSITY OF BIRMINGHAM

Visualisation

Week 1

What is Visualisation?

What Could this Mean?

I

x_1	y_1
10	8.04
8	6.95
13	7.58
9	8.81
11	8.33
14	9.96
6	7.24
4	4.26
12	10.84
7	4.82
5	5.68

II

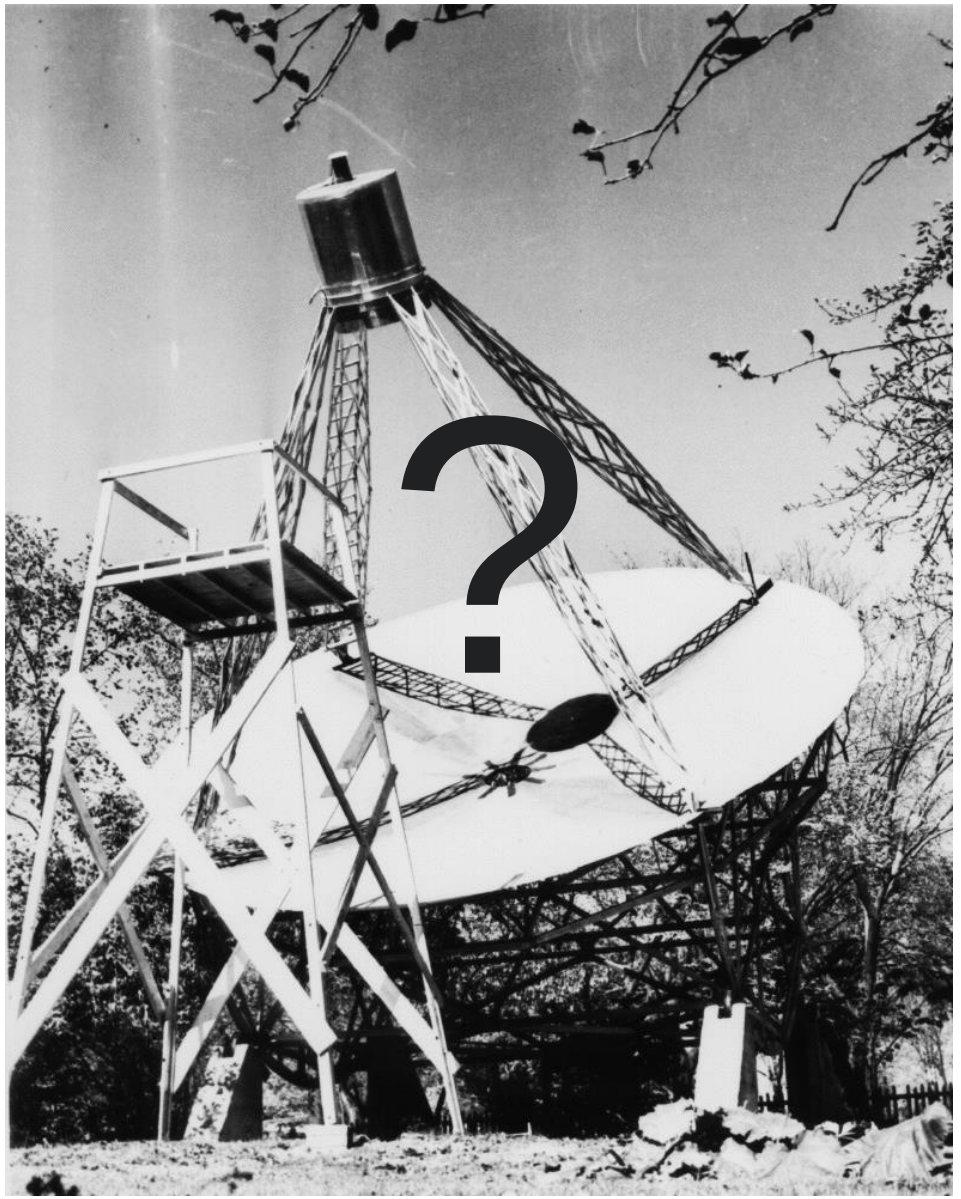
x_2	y_2
10	9.14
8	8.14
13	8.74
9	8.77
11	9.26
14	8.1
6	6.13
4	3.1
12	9.13
7	7.26
5	4.74

III

x_3	y_3
10	7.46
8	6.77
13	12.74
9	7.11
11	7.81
14	8.84
6	6.08
4	5.39
12	8.15
7	6.42
5	5.73

IV

x_4	y_4
8	6.58
8	5.76
8	7.71
8	8.84
8	8.47
8	7.04
8	5.25
19	12.5
8	5.56
8	7.91
8	6.89



I		II		III		IV	
x_1	y_1	x_2	y_2	x_3	y_3	x_4	y_4
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

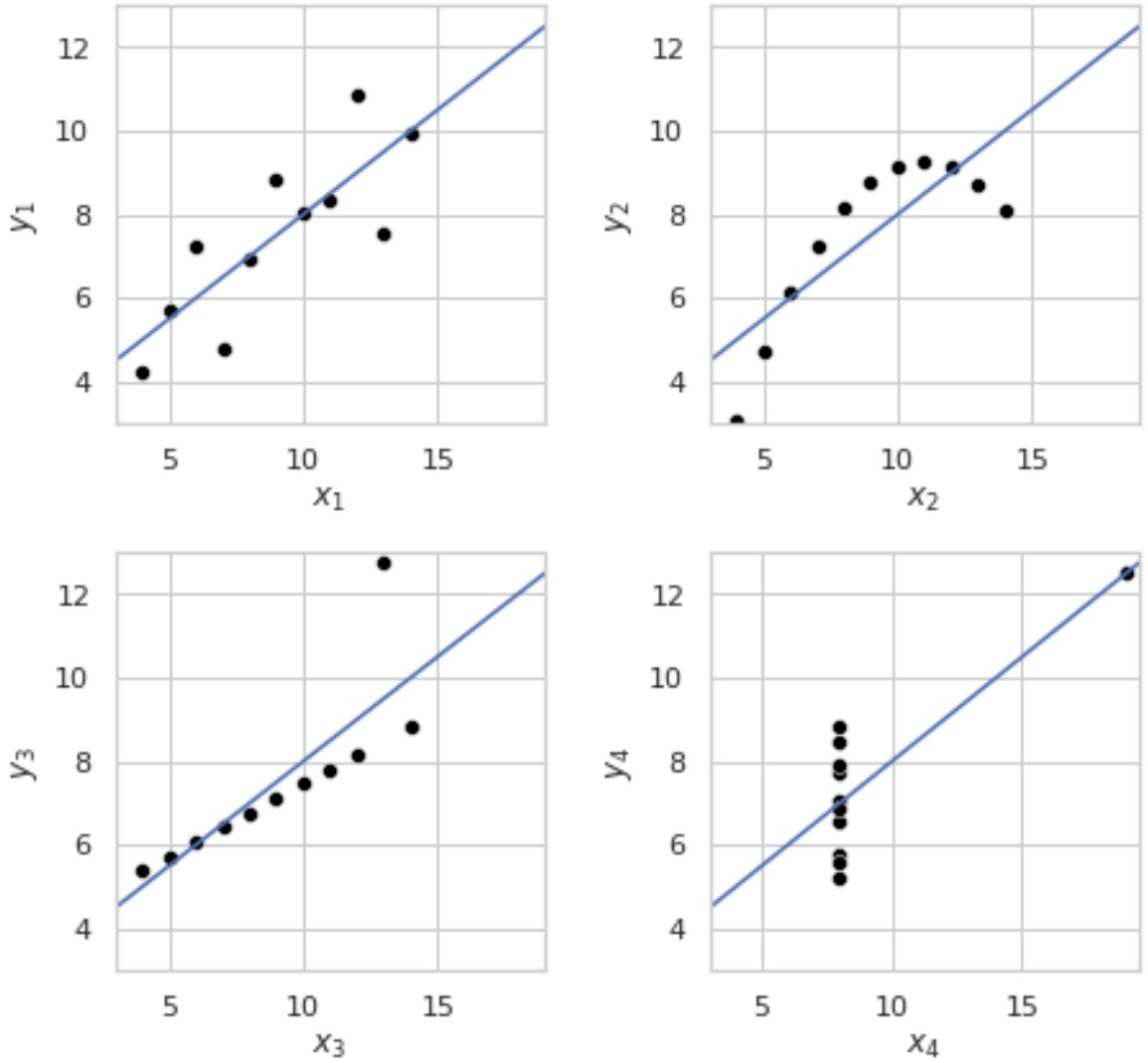
Mean:	9	7.5	9	7.5	9	7.5	9	7.5
Variance:	11	4.13	11	4.13	11	4.13	11	4.13
Correlation:	0.816		0.816		0.816		0.816	

Least squares regression $y = ax + b$

a:	0.5		0.5		0.5		0.5	
b:	0.3		0.3		0.3		0.3	

Anscombe Quartet

Frank Anscombe

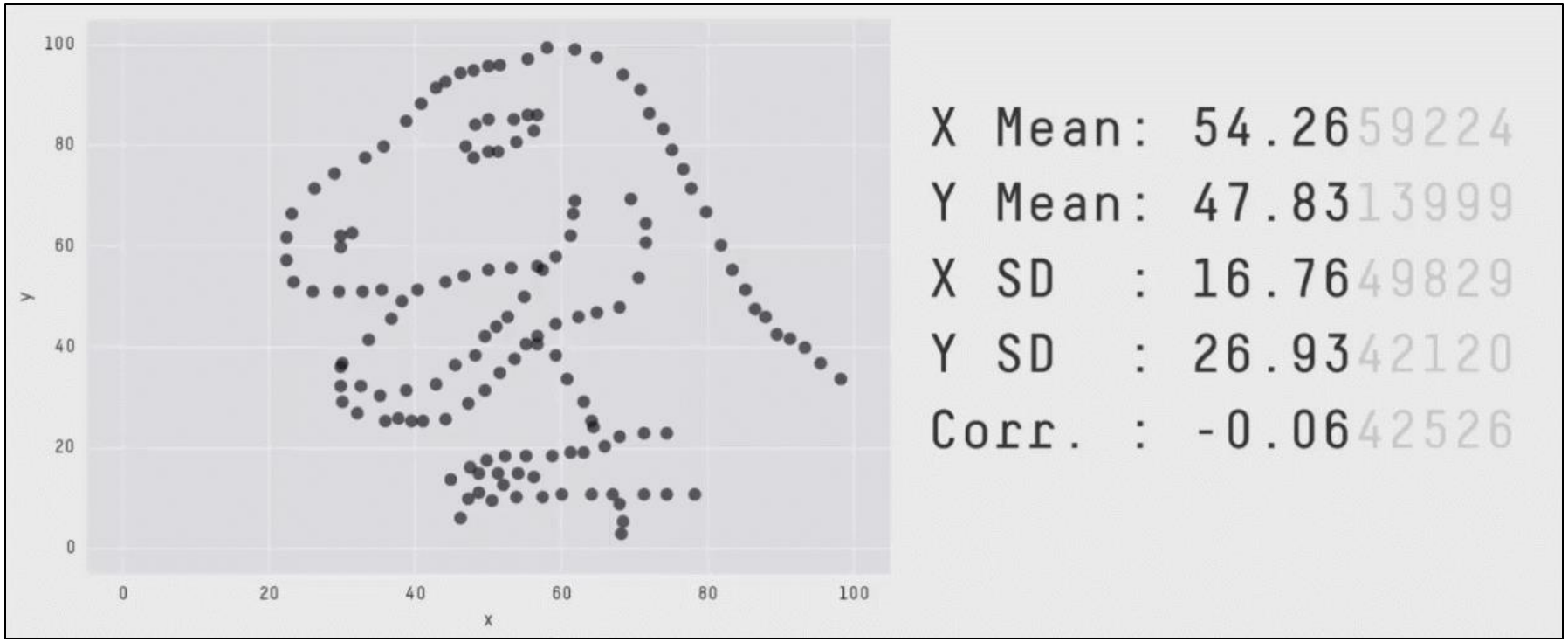


Anscombe (1973). Graphs in Statistical Analysis

Lesson:

- **Always look at your data**

Anscombosaurus



Alberto Cairo

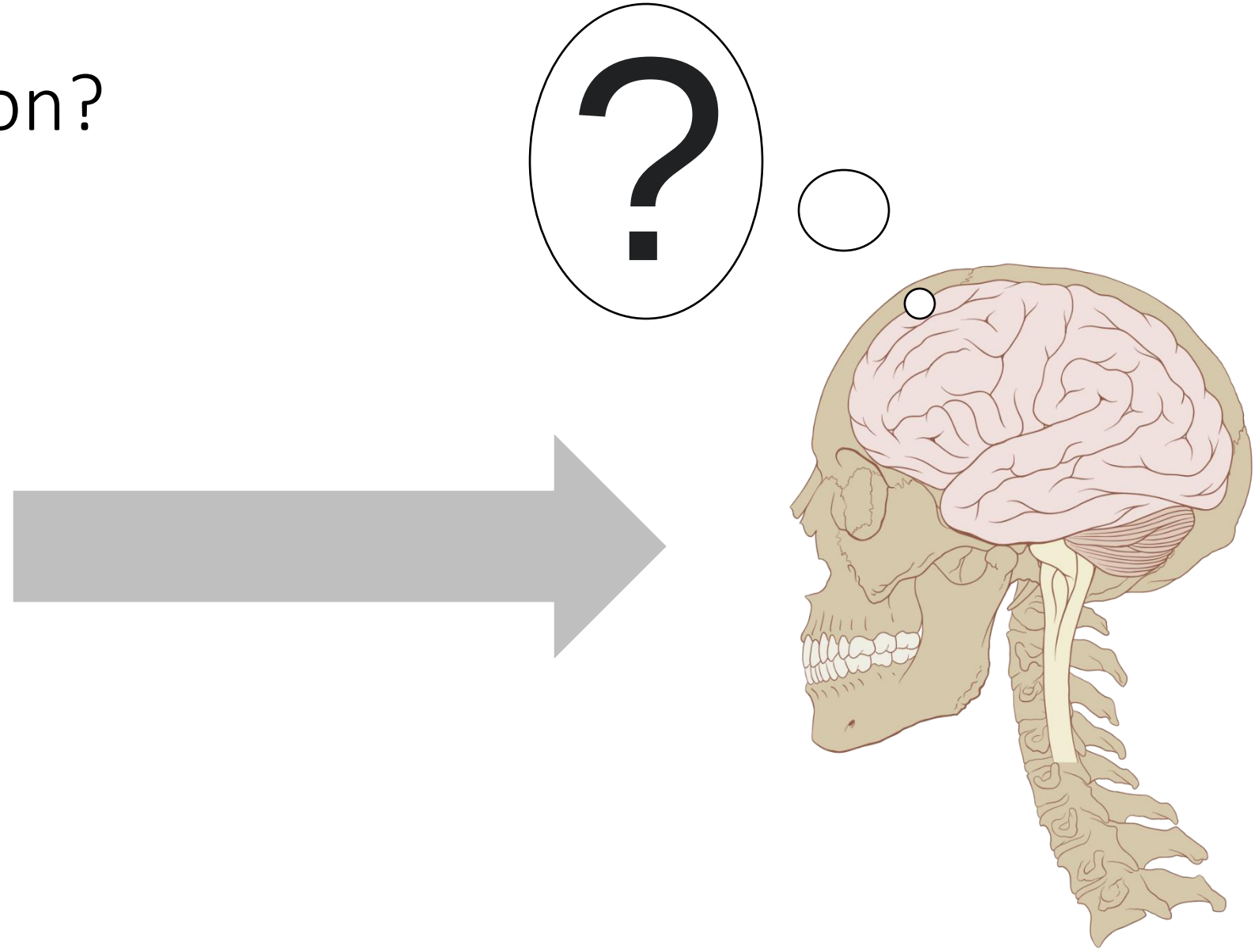
What is going on?

x_1	y_1
10	8.04
8	6.95
13	7.58
9	8.81
11	8.33
14	9.96
6	7.24
4	4.26
12	10.84
7	4.82
5	5.68

x_2	y_2
10	9.14
8	8.14
13	8.74
9	8.77
11	9.26
14	8.1
6	6.13
4	3.1
12	9.13
7	7.26
5	4.74

x_3	y_3
10	7.46
8	6.77
13	12.74
9	7.11
11	7.81
14	8.84
6	6.08
4	5.39
12	8.15
7	6.42
5	5.73

x_4	y_4
8	6.58
8	5.76
8	7.71
8	8.84
8	8.47
8	7.04
8	5.25
19	12.5
8	5.56
8	7.91
8	6.89



What is going on?

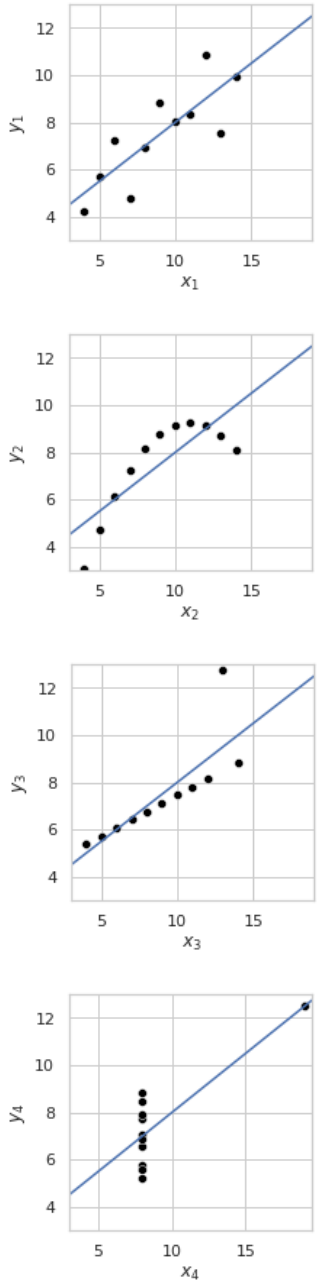
x_1	y_1
10	8.04
8	6.95
13	7.58
9	8.81
11	8.33
14	9.96
6	7.24
4	4.26
12	10.84
7	4.82
5	5.68

x_2	y_2
10	9.14
8	8.14
13	8.74
9	8.77
11	9.26
14	8.1
6	6.13
4	3.1
12	9.13
7	7.26
5	4.74

x_3	y_3
10	7.46
8	6.77
13	12.74
9	7.11
11	7.81
14	8.84
6	6.08
4	5.39
12	8.15
7	6.42
5	5.73

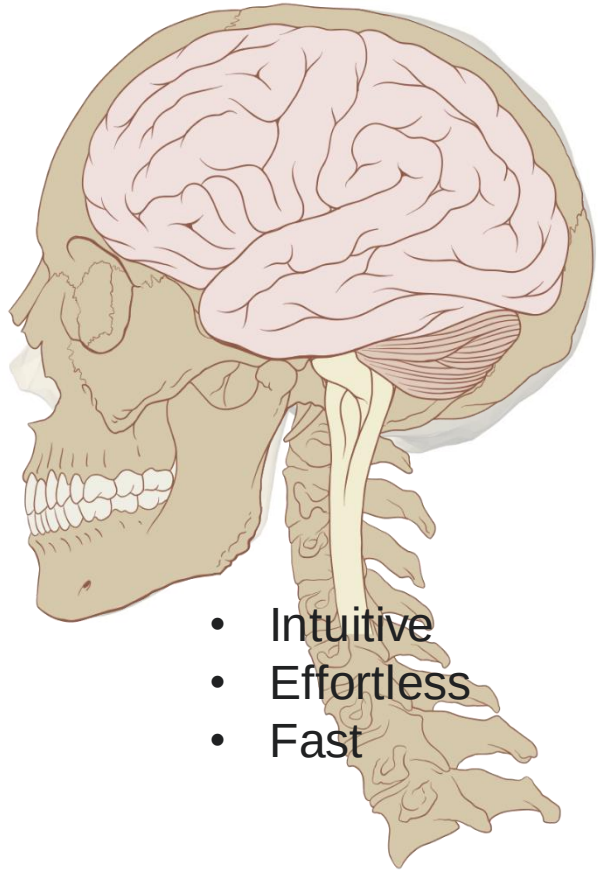
x_4	y_4
8	6.58
8	5.76
8	7.71
8	8.84
8	8.47
8	7.04
8	5.25
19	12.5
8	5.56
8	7.91
8	6.89

Encode



Visual cortex

An expert at pattern recognition

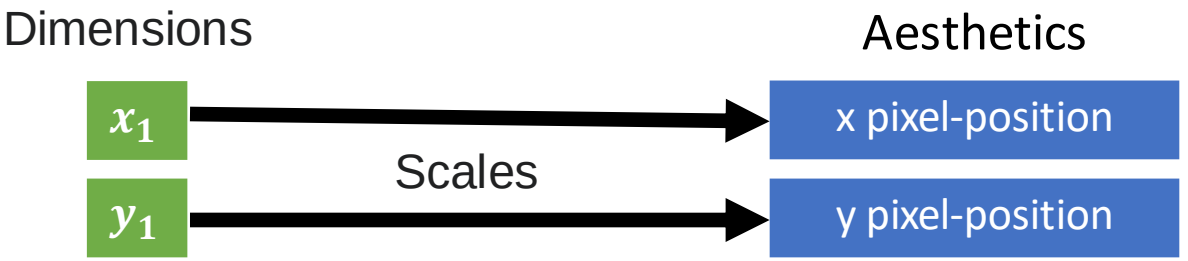


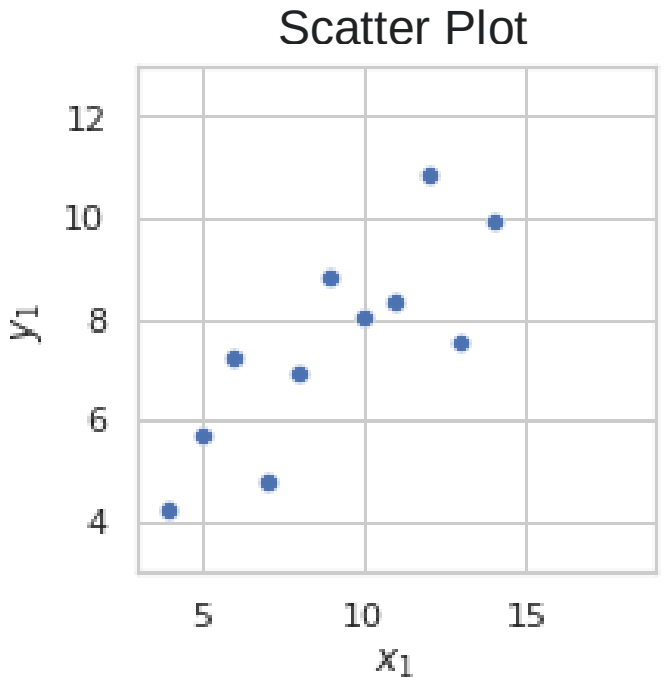
- Intuitive
- Effortless
- Fast

Mapping Variables to Aesthetics

Data

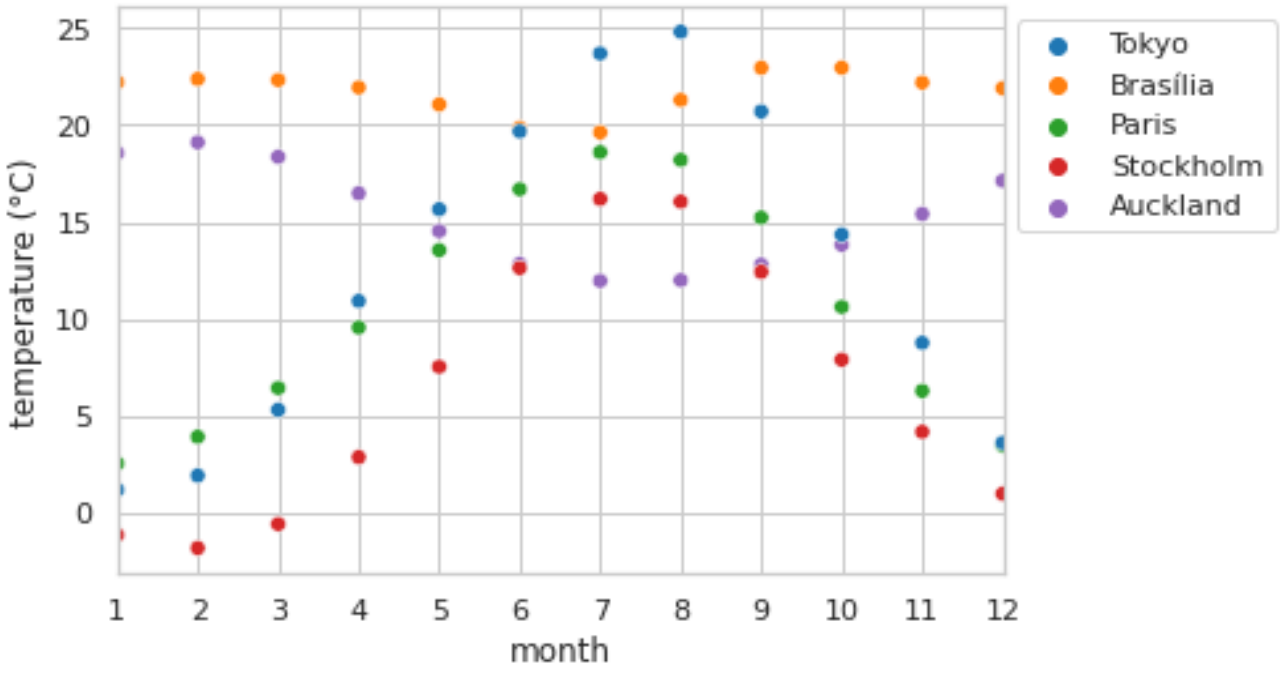
x_1	y_1
10	8.04
8	6.95
13	7.58
9	8.81
11	8.33
14	9.96
6	7.24
4	4.26
12	10.84
7	4.82
5	5.68

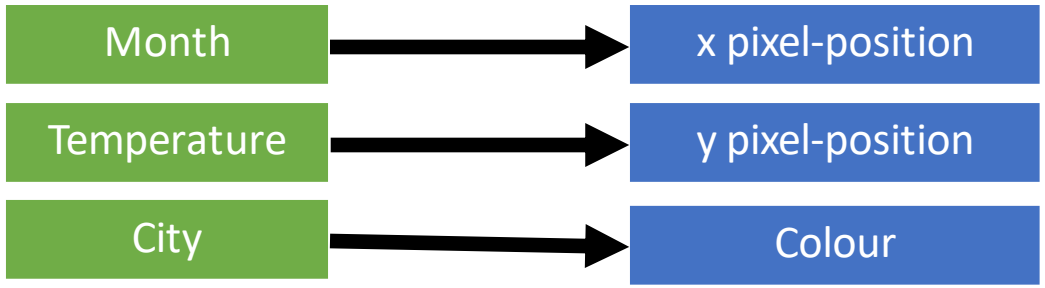




Different Ways to Map Data

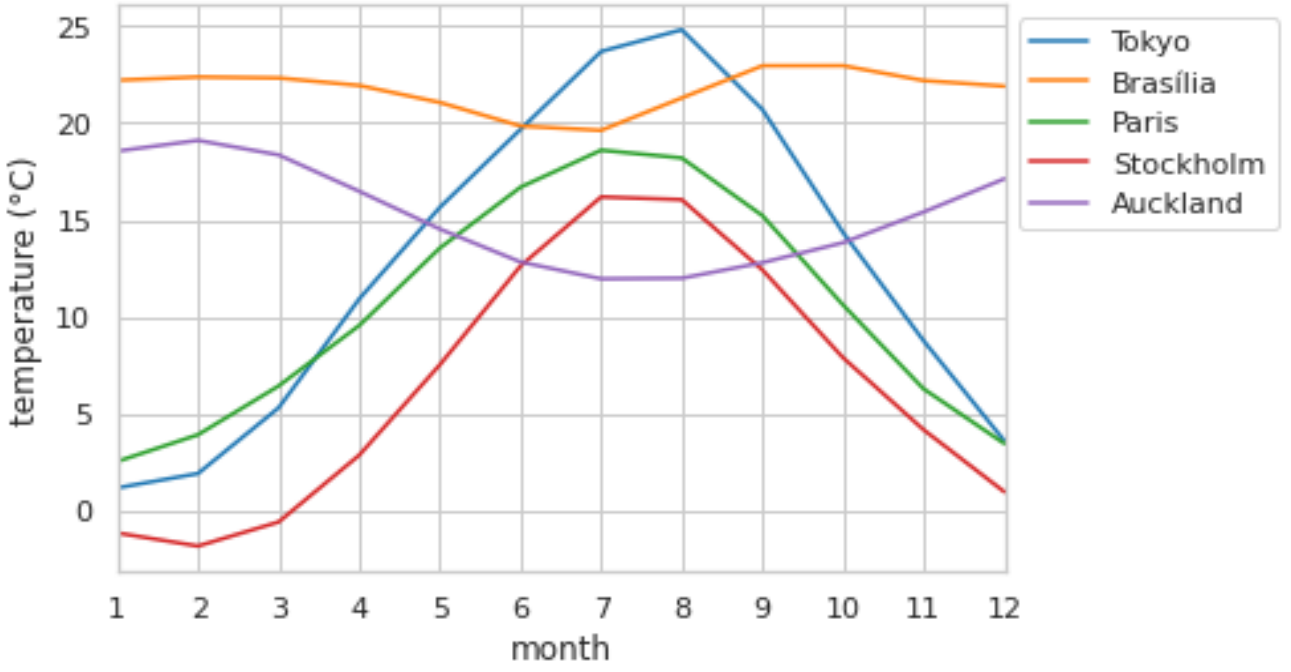
	month	City	temperature (C)
0	1	Auckland	18.551830
1	1	Brasília	22.216207
2	1	Paris	2.513351
3	1	Stockholm	-1.182832
4	1	Tokyo	1.148084
5	2	Auckland	19.106219
6	2	Brasília	22.382339
7	2	Paris	3.892574
8	2	Stockholm	-1.858202
9	2	Tokyo	1.888856
10	3	Auckland	18.357038
11	3	Brasília	22.332715
12	3	Paris	6.413973
13	3	Stockholm	-0.616388
14	3	Tokyo	5.285594

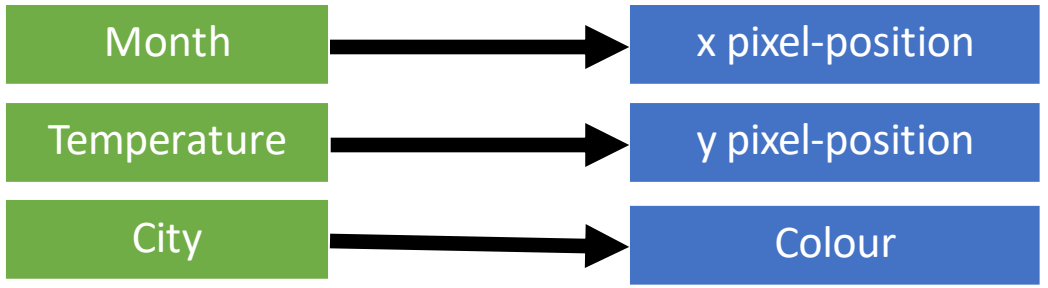




Different Ways to Map Data

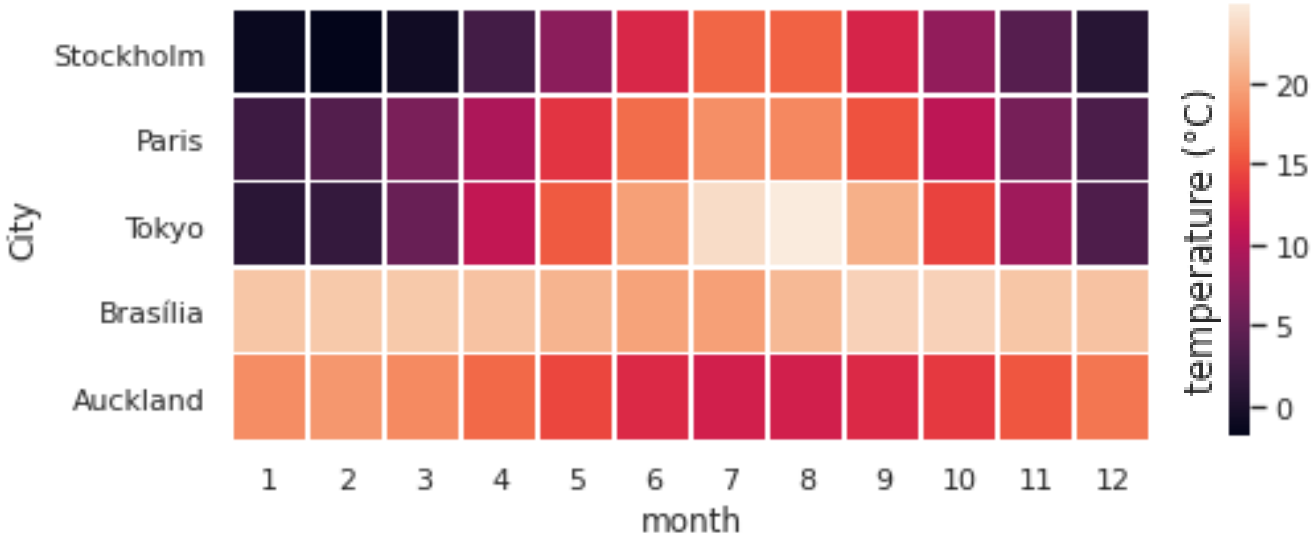
	month	City	temperature (C)
0	1	Auckland	18.551830
1	1	Brasília	22.216207
2	1	Paris	2.513351
3	1	Stockholm	-1.182832
4	1	Tokyo	1.148084
5	2	Auckland	19.106219
6	2	Brasília	22.382339
7	2	Paris	3.892574
8	2	Stockholm	-1.858202
9	2	Tokyo	1.888856
10	3	Auckland	18.357038
11	3	Brasília	22.332715
12	3	Paris	6.413973
13	3	Stockholm	-0.616388
14	3	Tokyo	5.285594

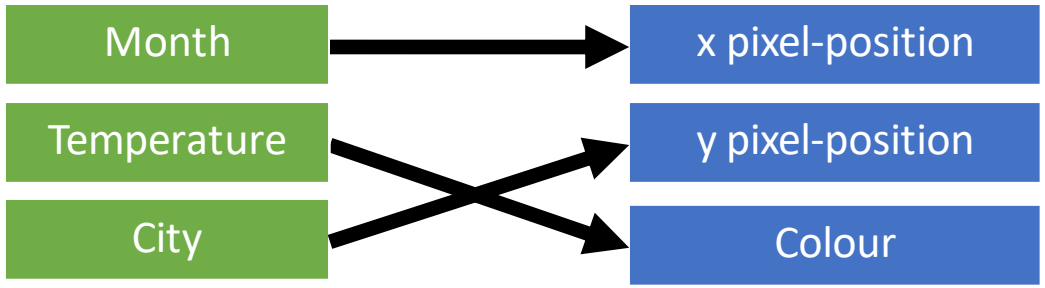




Different Ways to Map Data

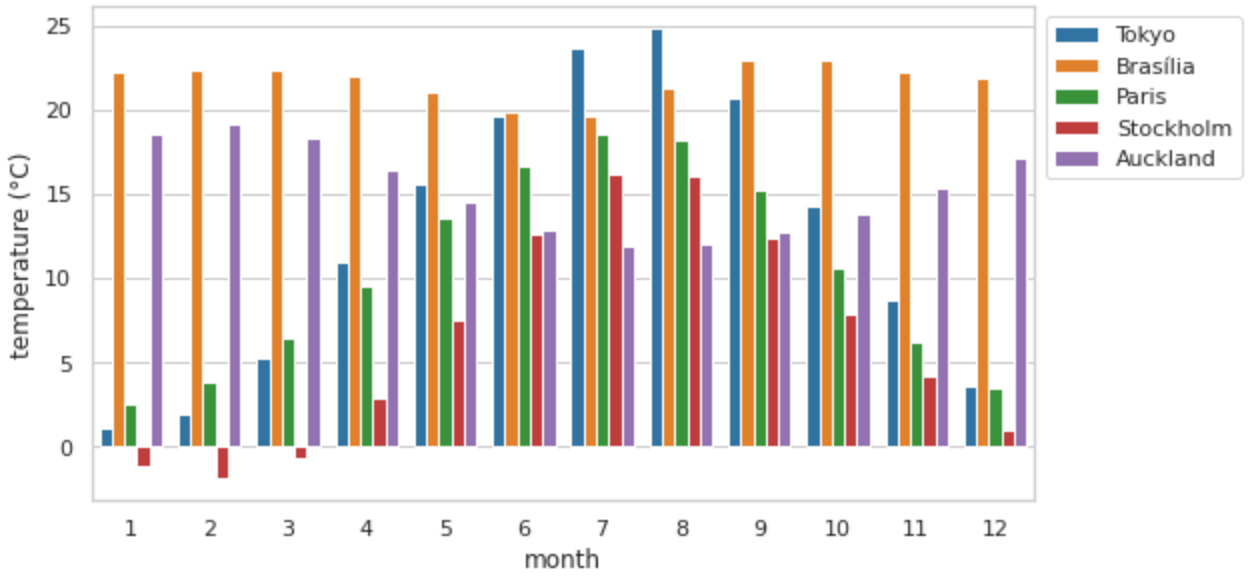
	month	City	temperature (C)
0	1	Auckland	18.551830
1	1	Brasília	22.216207
2	1	Paris	2.513351
3	1	Stockholm	-1.182832
4	1	Tokyo	1.148084
5	2	Auckland	19.106219
6	2	Brasília	22.382339
7	2	Paris	3.892574
8	2	Stockholm	-1.858202
9	2	Tokyo	1.888856
10	3	Auckland	18.357038
11	3	Brasília	22.332715
12	3	Paris	6.413973
13	3	Stockholm	-0.616388
14	3	Tokyo	5.285594

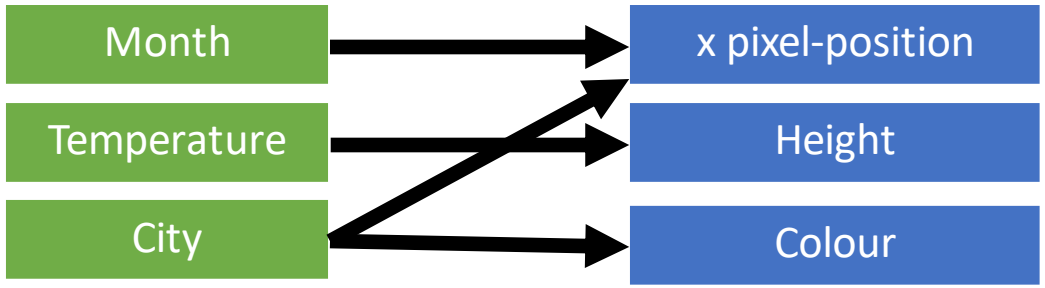


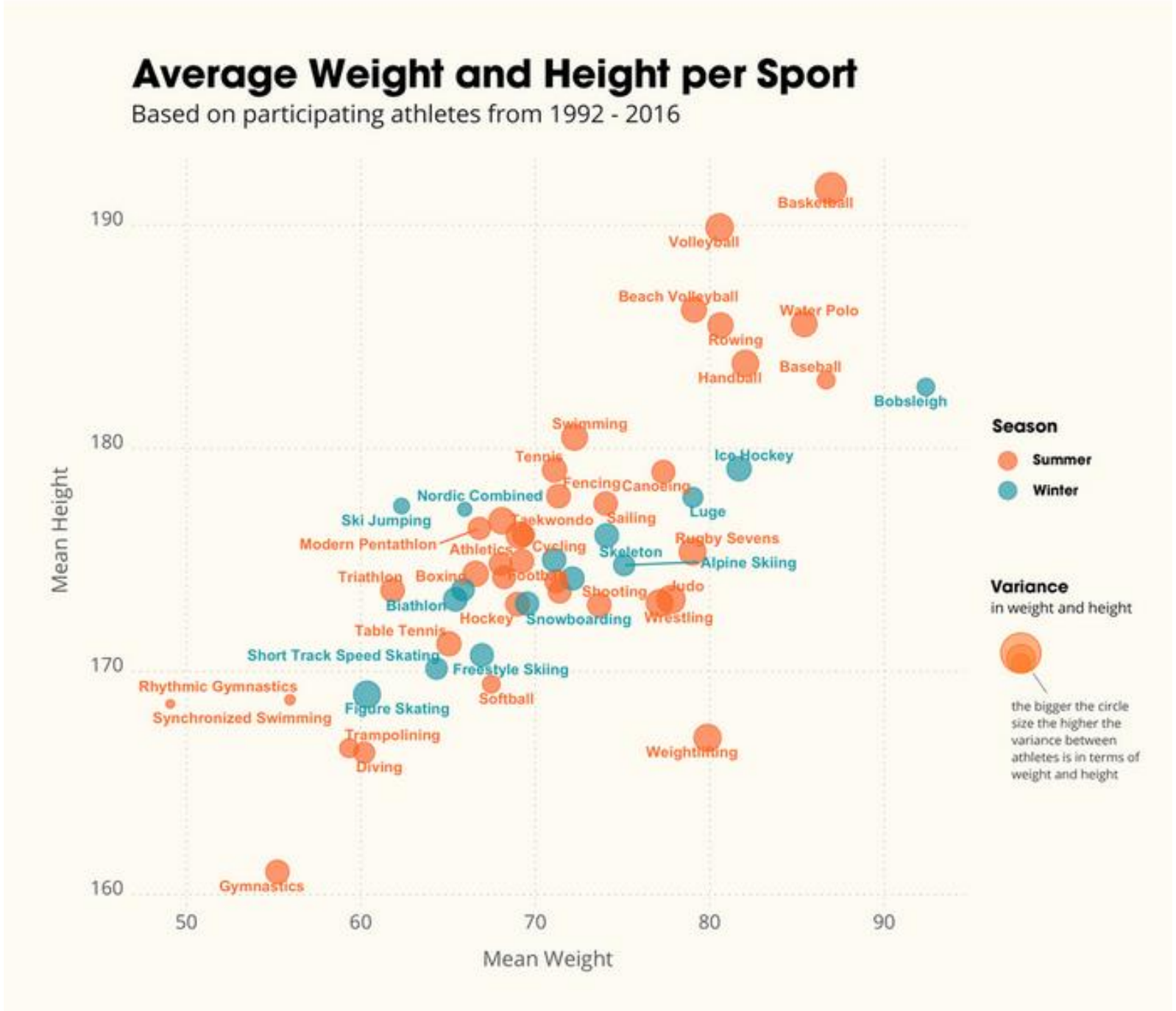


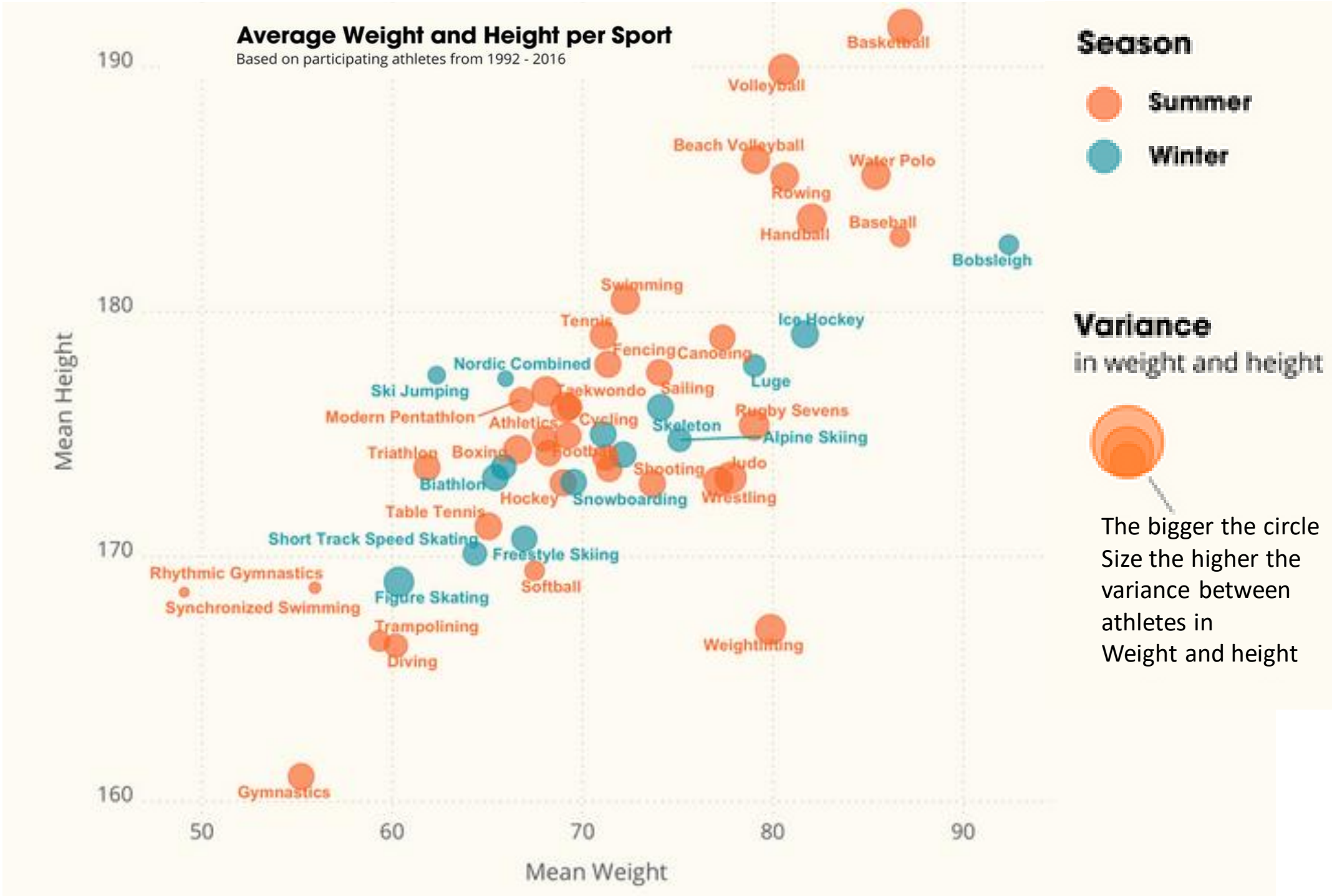
Different Ways to Map Data

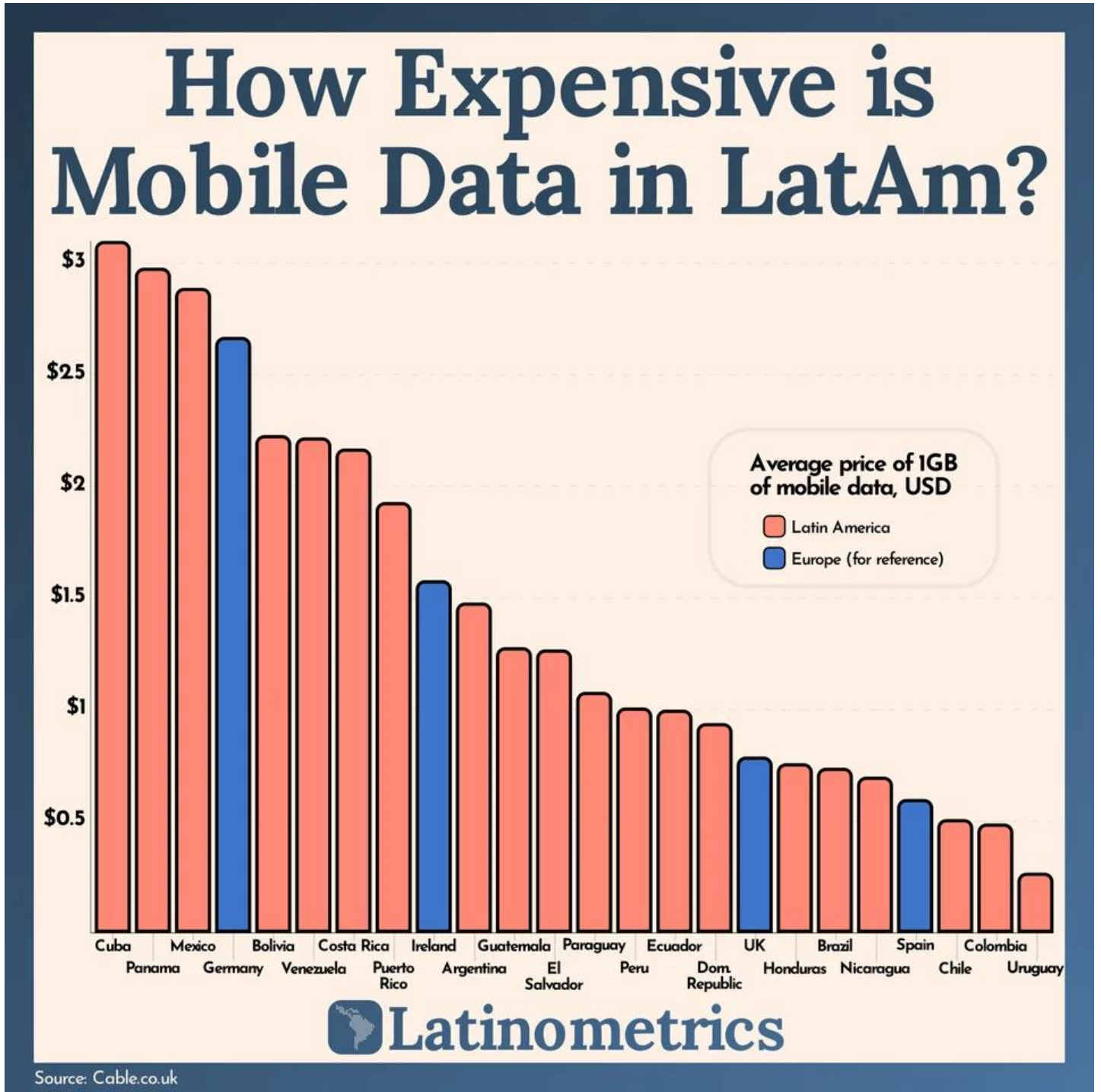
	month	City	temperature (C)
0	1	Auckland	18.551830
1	1	Brasília	22.216207
2	1	Paris	2.513351
3	1	Stockholm	-1.182832
4	1	Tokyo	1.148084
5	2	Auckland	19.106219
6	2	Brasília	22.382339
7	2	Paris	3.892574
8	2	Stockholm	-1.858202
9	2	Tokyo	1.888856
10	3	Auckland	18.357038
11	3	Brasília	22.332715
12	3	Paris	6.413973
13	3	Stockholm	-0.616388
14	3	Tokyo	5.285594



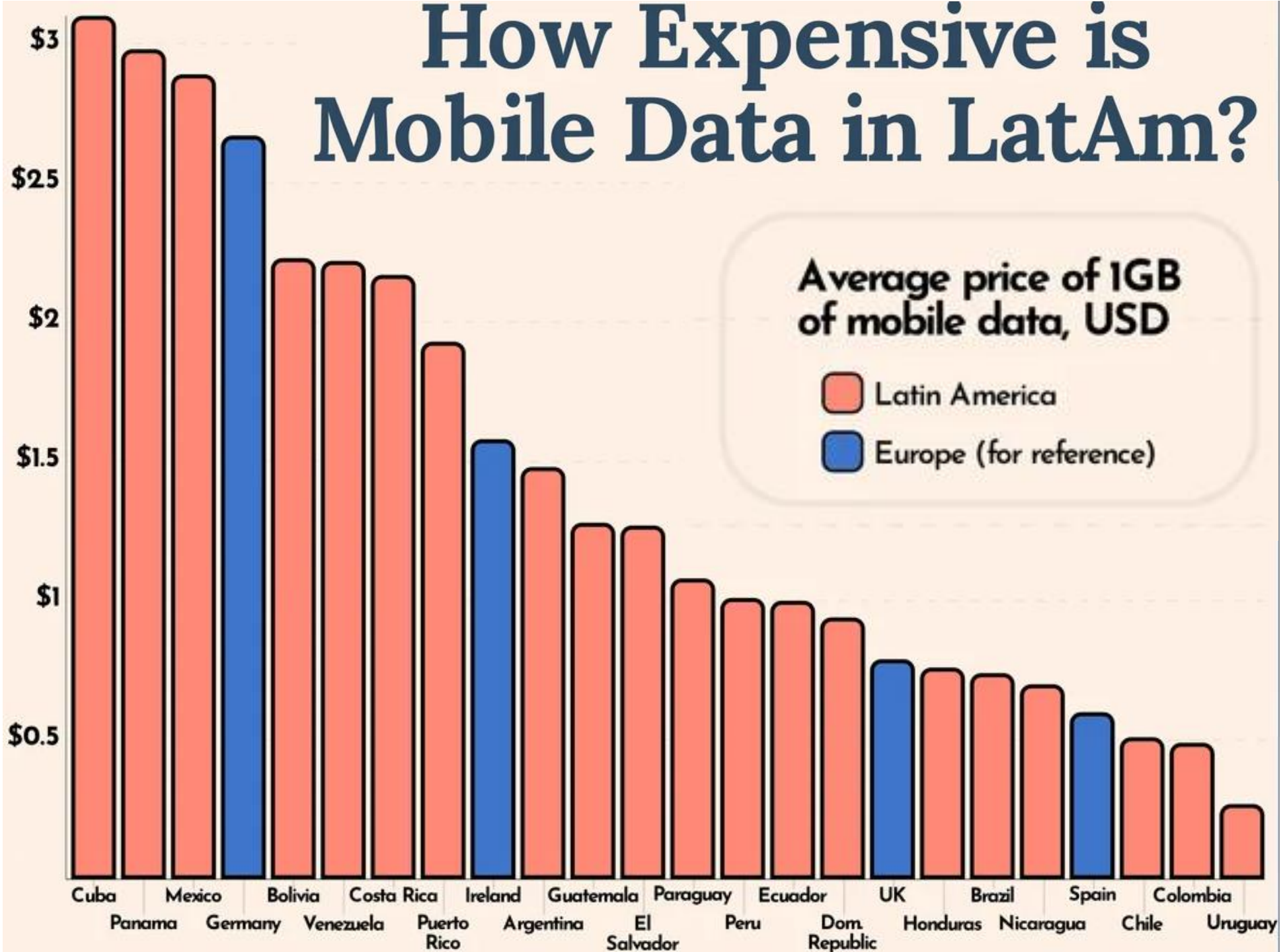



Average Weight and Height per Sport
Based on participating athletes from 1992 - 2016
Mean Height
190
180
170
160
50
60
70
80
90
Mean Weight
Basketball
Volleyball
Beach Volleyball
Water Polo
Rowing
Handball
Baseball
Bobsleigh
Swimming
Tennis
Ice Hockey
Fencing
Canoeing
Nordic Combined
Luge
Ski Jumping
Taekwondo
Sailing
Modern Pentathlon
Athletics
Cycling
Skeleton
Rugby Sevens
Alpine Skiing
Triathlon
Boxing
Shooting
Judo
Biathlon
Hockey
Snowboarding
Wrestling
Table Tennis
Short Track Speed Skating
Freestyle Skiing
Rhythmic Gymnastics
Softball
Synchronized Swimming
Figure Skating
Trampolining
Weightlifting
Diving
Gymnastics
Season
Summer
Winter
Variance
in weight and height
the bigger the circle size the higher the variance between athletes is in terms of weight and height


Average Weight and Height per Sport
Based on participating athletes from 1992 - 2016
Mean Height
Mean Weight
190
180
170
160
50
60
70
80
90
Basketball
Volleyball
Beach Volleyball
Water Polo
Rowing
Handball
Baseball
Bobsleigh
Swimming
Tennis
Ice Hockey
Fencing
Canoeing
Nordic Combined
Luge
Ski Jumping
Taekwondo
Sailing
Modern Pentathlon
Athletics
Cycling
Skeleton
Rugby Sevens
Alpine Skiing
Triathlon
Boxing
Shooting
Judo
Biathlon
Hockey
Snowboarding
Wrestling
Table Tennis
Short Track Speed Skating
Freestyle Skiing
Rhythmic Gymnastics
Softball
Synchronized Swimming
Figure Skating
Trampolining
Diving
Weightlifting
Gymnastics
Season
Summer
Winter
Variance
in weight and height
The bigger the circle
Size the higher the
variance between
athletes in
Weight and height

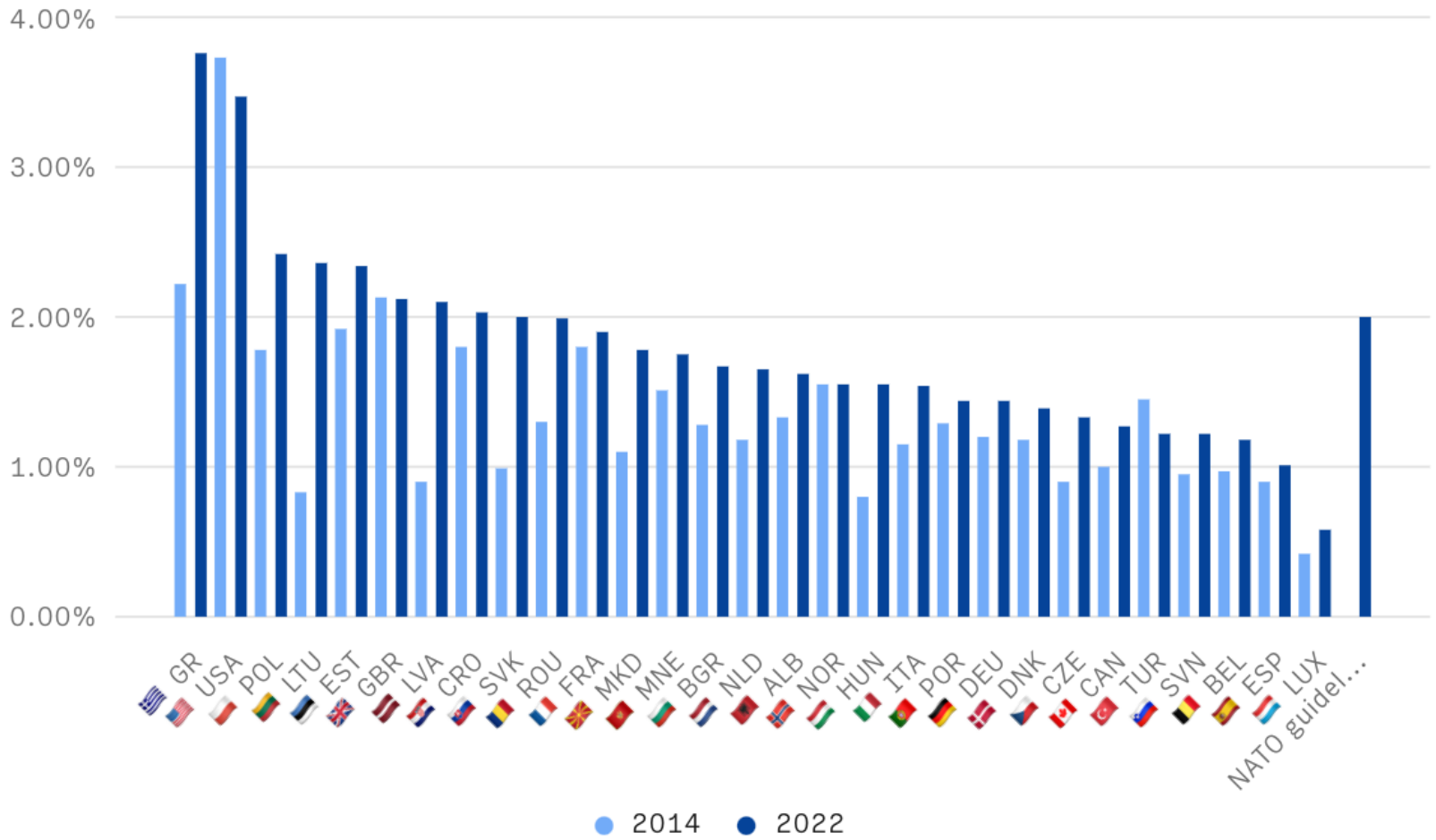

How Expensive is Mobile Data in LatAm?
$3
$2.5
$2
$1.5
$1
$0.5
Average price of 1GB of mobile data, USD
Latin America
Europe (for reference)
Cuba
Panama
Mexico
Germany
Bolivia
Venezuela
Costa Rica
Puerto Rico
Ireland
Argentina
Guatemala
El Salvador
Paraguay
Peru
Ecuador
Dom. Republic
UK
Honduras
Brazil
Nicaragua
Spain
Chile
Colombia
Uruguay
Latinometrics
Source: Cable.co.uk

How Expensive is Mobile Data in LatAm?



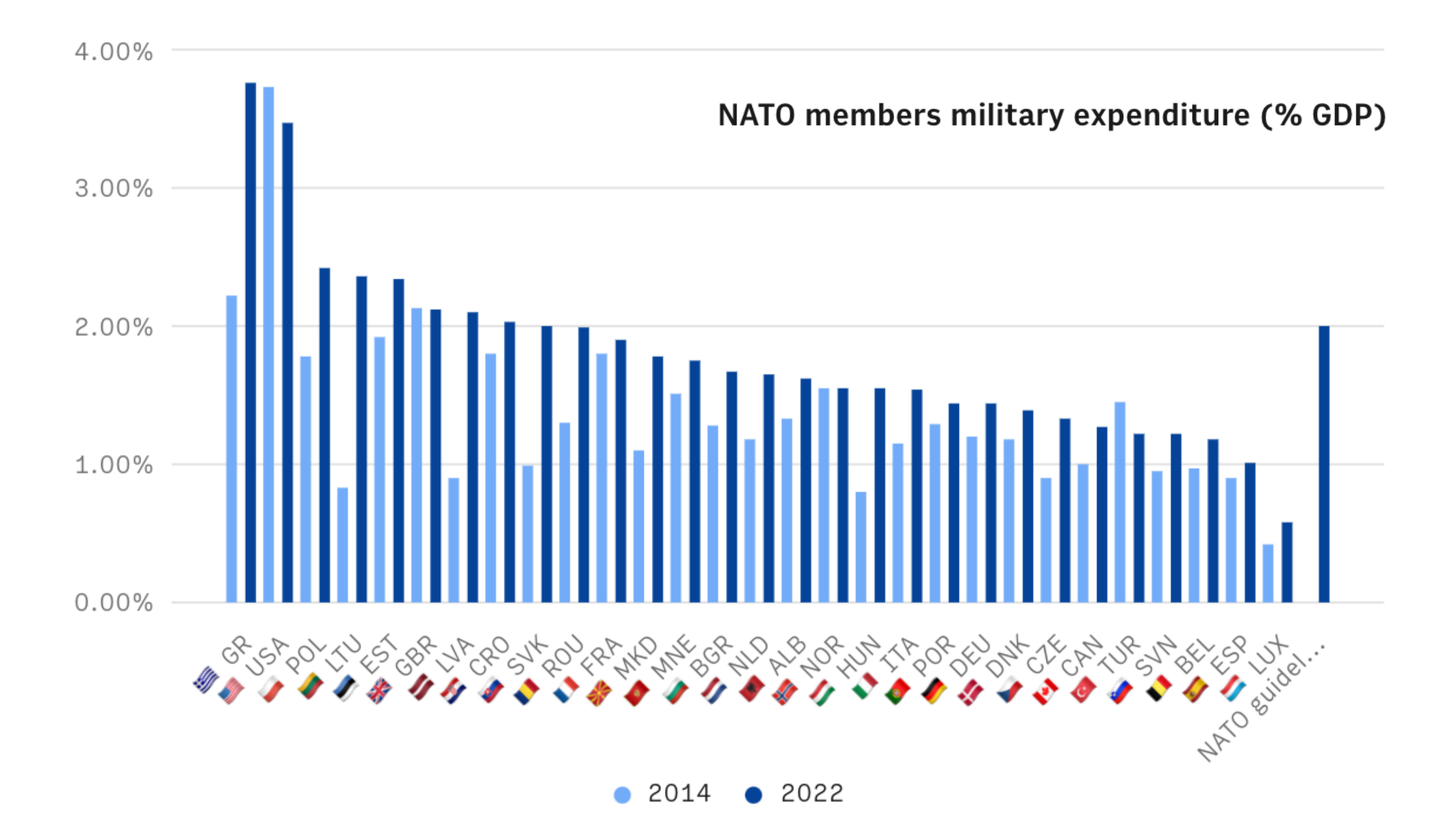
NATO members military expenditure (% GDP)

Military expenditure as % of GDP, 2014 vs 2022. Based on 2015 price and exchange rates.



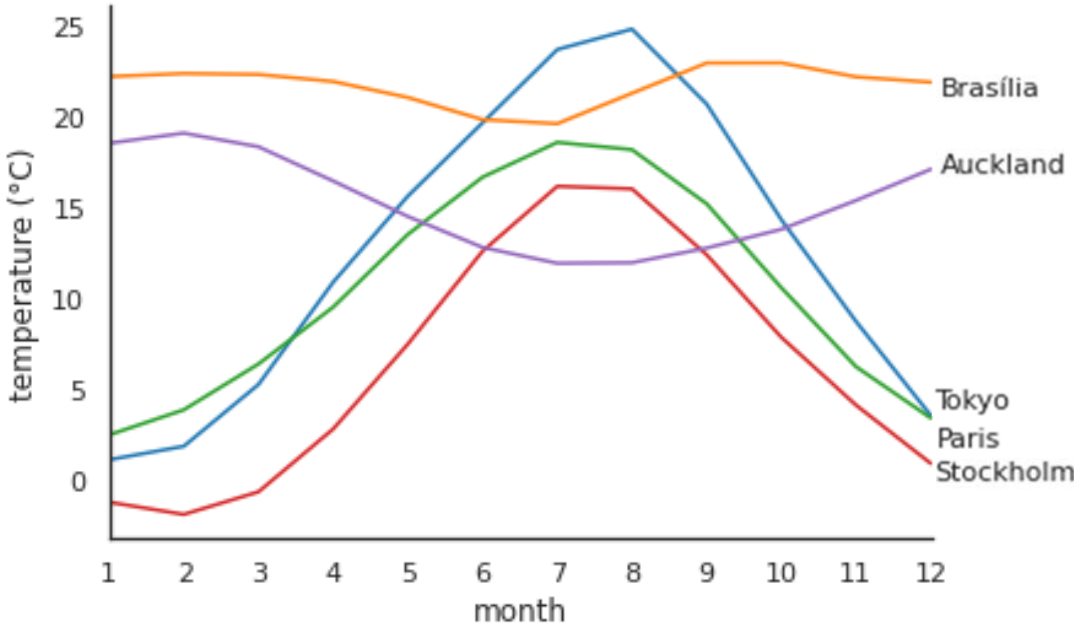
Source: https://www.nato.int/nato_static_fl2014/assets/pdf/2022/6/pdf/220627-def-exp-2022-en.pdf

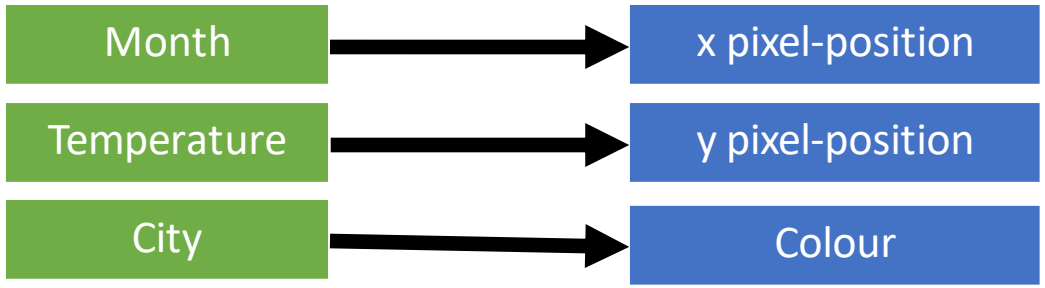
Created by **rows.com/alberto_m**


NATO members military expenditure (% GDP)
4.00%
3.00%
2.00%
1.00%
0.00%
GR
USA
POL
LTU
EST
GBR
LVA
CRO
SVK
ROU
FRA
MKD
MNE
BGR
NLD
ALB
NOR
HUN
ITA
POR
DEU
DNK
CZE
CAN
TUR
SVN
BEL
ESP
LUX
NATO guidel...
2014
2022

Different Ways to Map Data

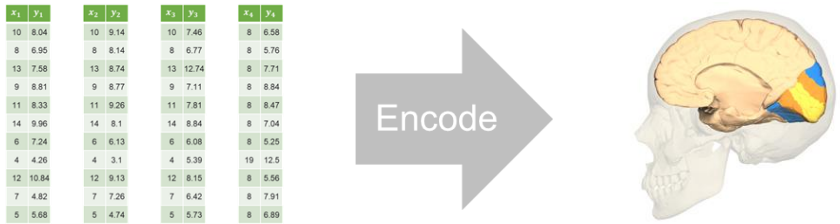
	month	City	temperature (C)
0	1	Auckland	18.551830
1	1	Brasília	22.216207
2	1	Paris	2.513351
3	1	Stockholm	-1.182832
4	1	Tokyo	1.148084
5	2	Auckland	19.106219
6	2	Brasília	22.382339
7	2	Paris	3.892574
8	2	Stockholm	-1.858202
9	2	Tokyo	1.888856
10	3	Auckland	18.357038
11	3	Brasília	22.332715
12	3	Paris	6.413973
13	3	Stockholm	-0.616388
14	3	Tokyo	5.285594



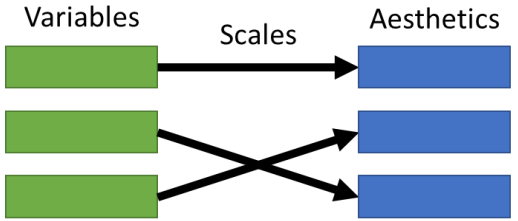


Summary - What is Data Visualisation?

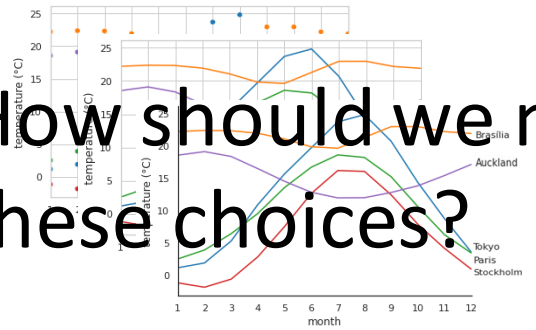
- Translate data for our visual brain.
  - Fast, intuitive processing



- Mapping variables to aesthetics.



- We can choose:
  - Which type of chart?
  - Which mapping?
  - Grid, legend, decoration?
  - …



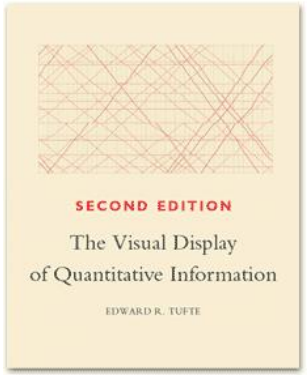
How should we make these choices?

What can Visualization Achieve?

Exploration:

- What is going on in a dataset?
- We do not know what we are looking for.
- Look at the data in different ways.
- Discover patterns.

Nathan Yau (2013)

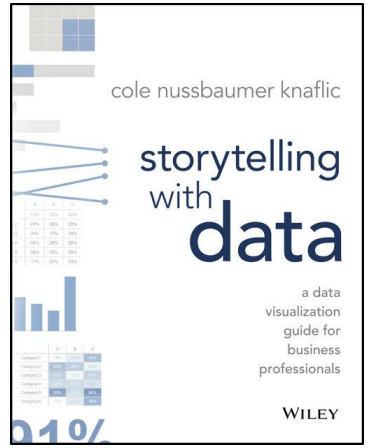


Edward Tufte (2001)

Explanation:

- We have a particular insight or message.
- Visualization is a communication tool.

- Highlight what is important.
- Remove irrelevant clutter.
- Show relevant context.
- Make it easily digestible.



Nussbaumer Knaflic (2015)

UNIVERSITY OF BIRMINGHAM

Visualisation

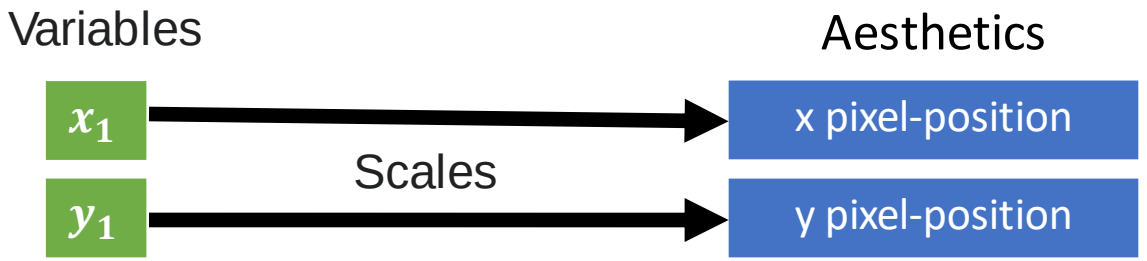
Week 1

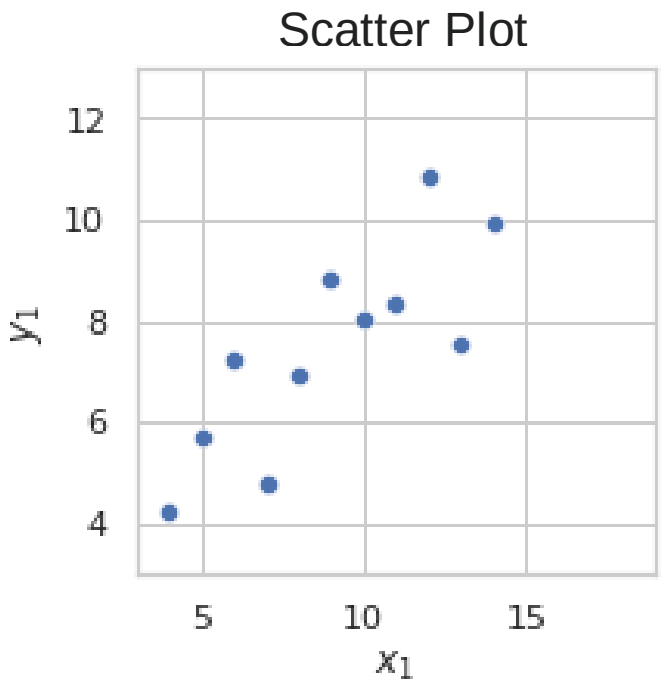
Variables and Aesthetics

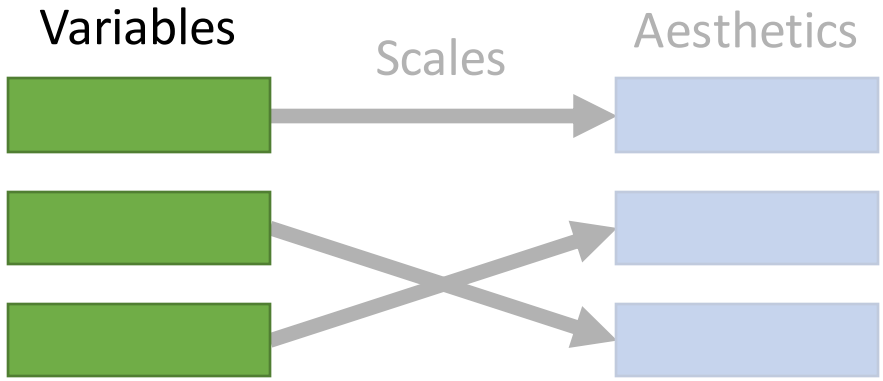
Mapping Variables to Aesthetics

Data

x_1	y_1
10	8.04
8	6.95
13	7.58
9	8.81
11	8.33
14	9.96
6	7.24
4	4.26
12	10.84
7	4.82
5	5.68







Numerical

aka
Quantitative

Continuous:

10.14, 14.1, 11.0, …
(e.g. temperatures)

Discrete:

1,2,3,… (indexes)
10, 13, 37, 4,… (counts)
1998, 1999, 2000 (years)

0.0, 0.25, 0.5, 0.75, 1.0
(real numbers with only discrete values allowed)

Categorical

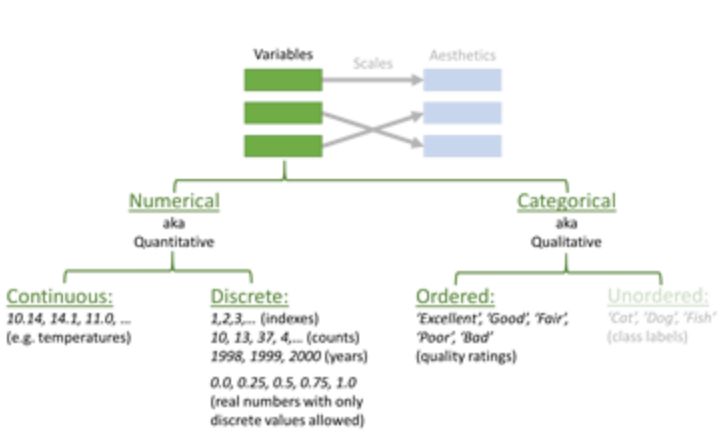
aka
Qualitative

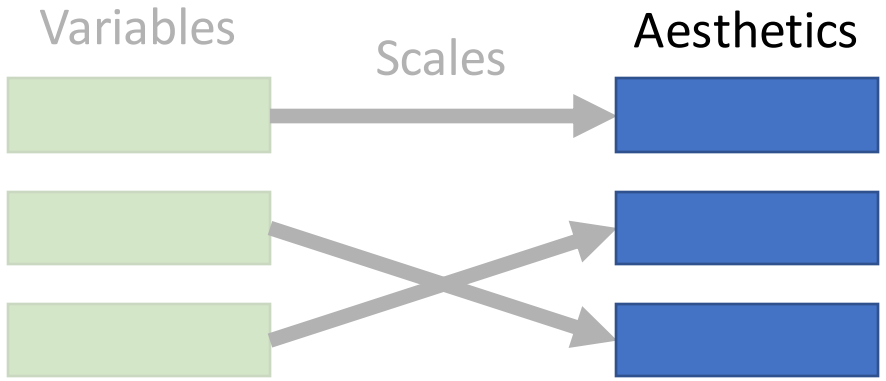
Ordered:

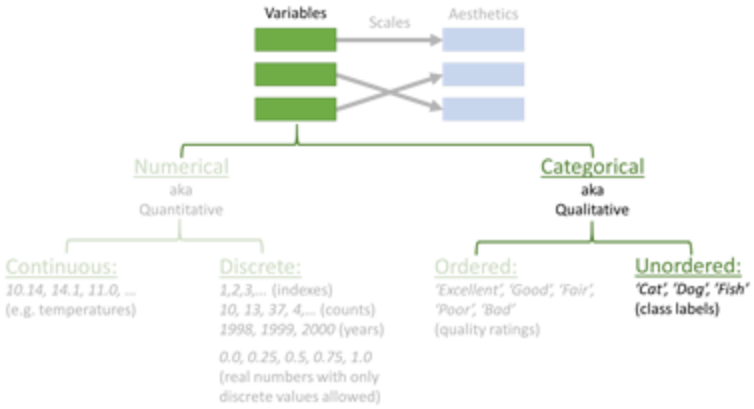
'Excellent', 'Good', 'Fair', 'Poor', 'Bad'
(quality ratings)

Unordered:

'Cat', 'Dog', 'Fish'
(class labels)

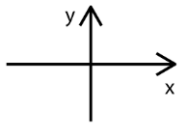






Ordered

(for **numerical** or **ordered categorical** variables)

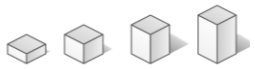


Position

Size (area)

Tilt

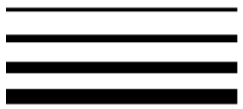
Length



3D volume

Curvature

Colour

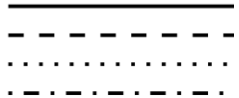


Line width

Unordered

(for **unordered categorical** variables)

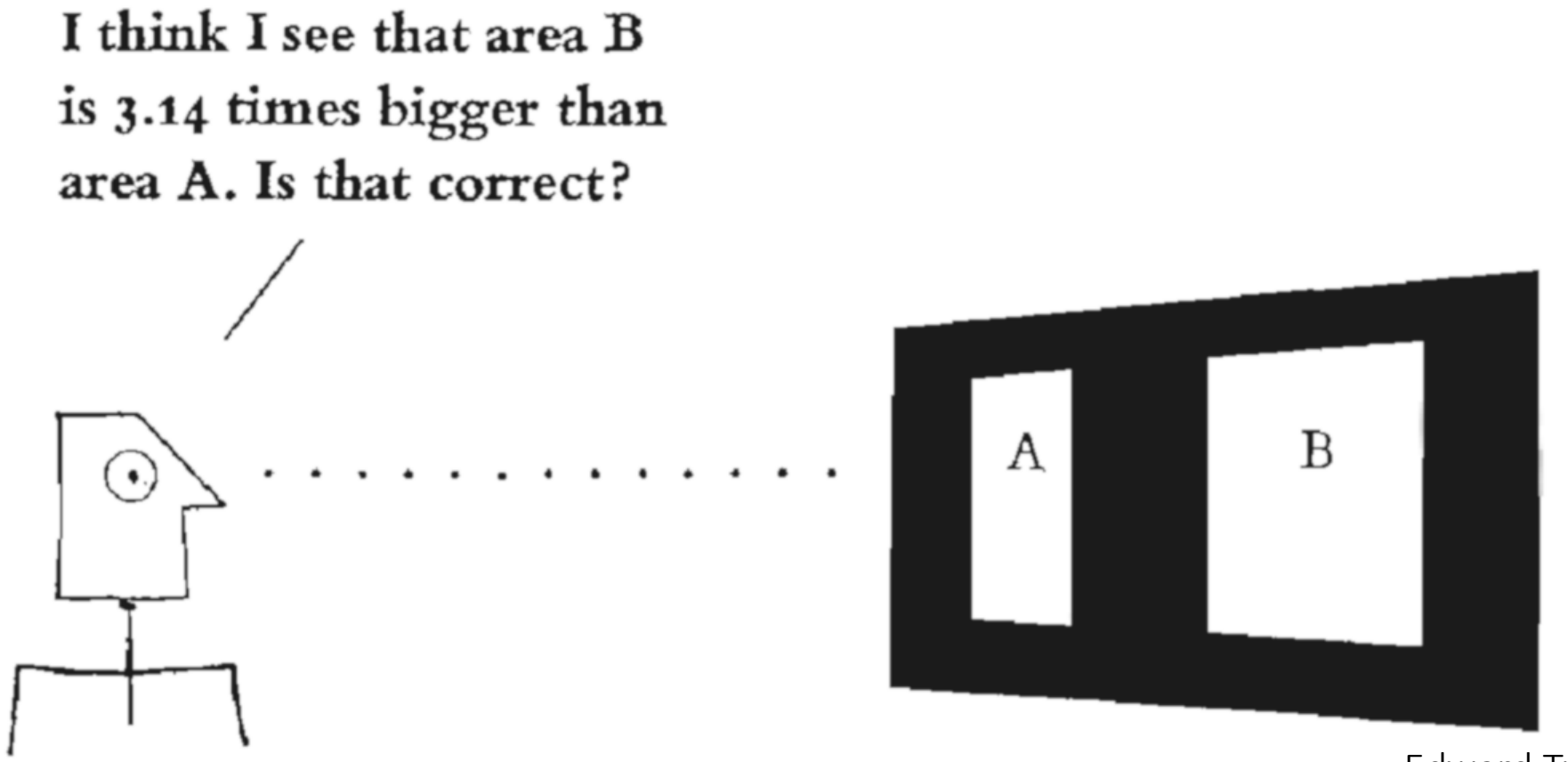
Shape



Line style

Colour

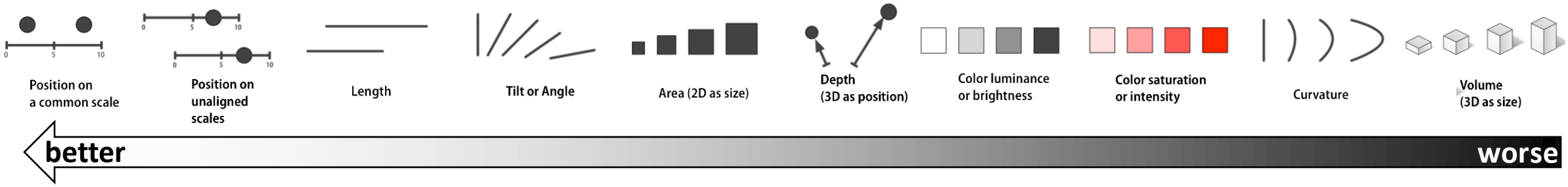
Are Some Aesthetics Better than Others?



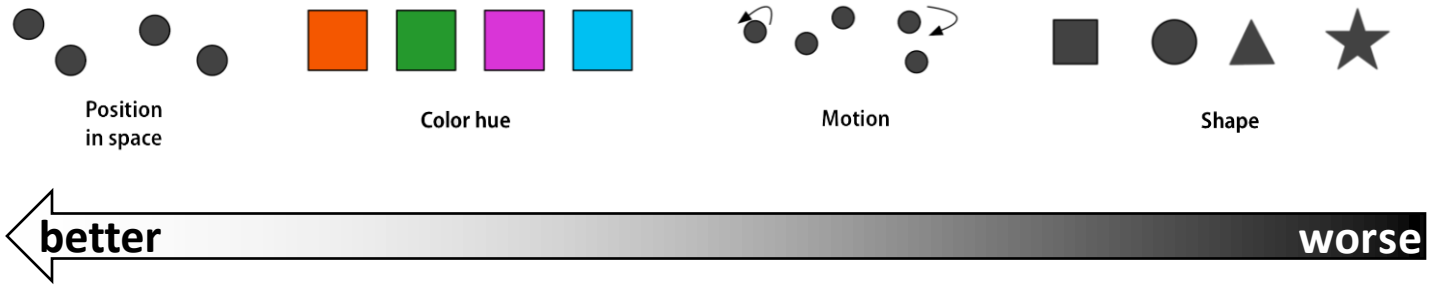
Edward Tufte

Are some aesthetics better than others?

Ordered aesthetics:



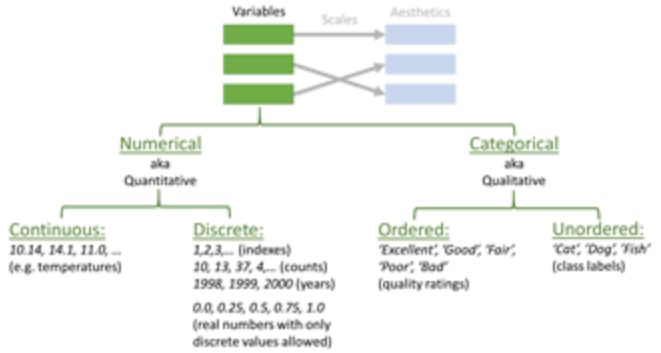
Unordered aesthetics:



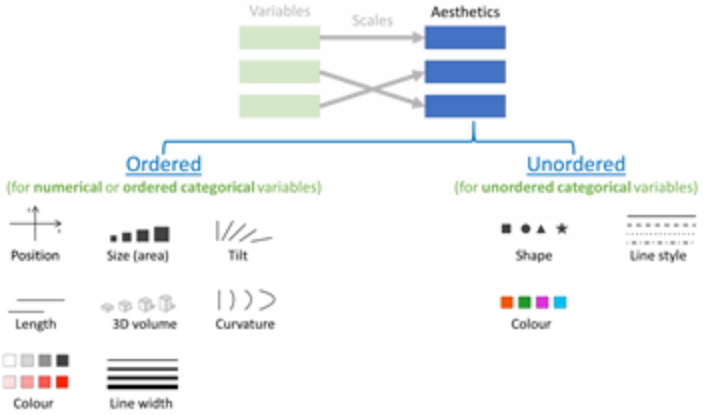
Kieran Healy. Data Visualization an Introduction (2018), Tamara Munzer (2015)

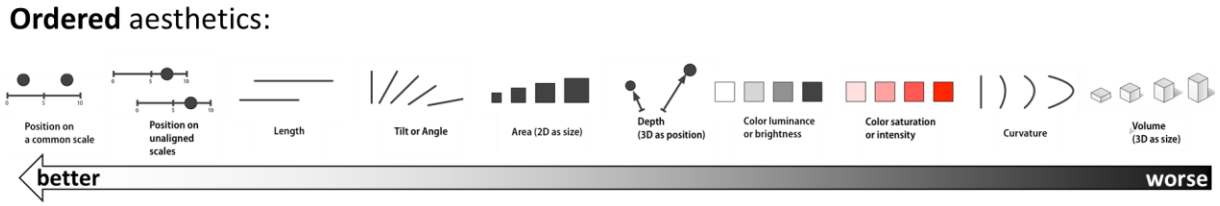
Summary

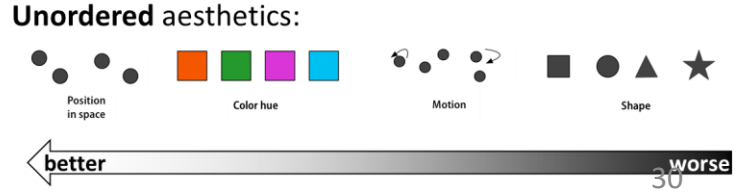
- Variables can be categorised.
  - Numerical, categorical
  - Continuous, discrete
  - Ordered, unordered



- Aesthetics are not created equal.
  - Ordered, unordered
  - Position and length most accurate.
  - Area and volume not so good.







What's Next?