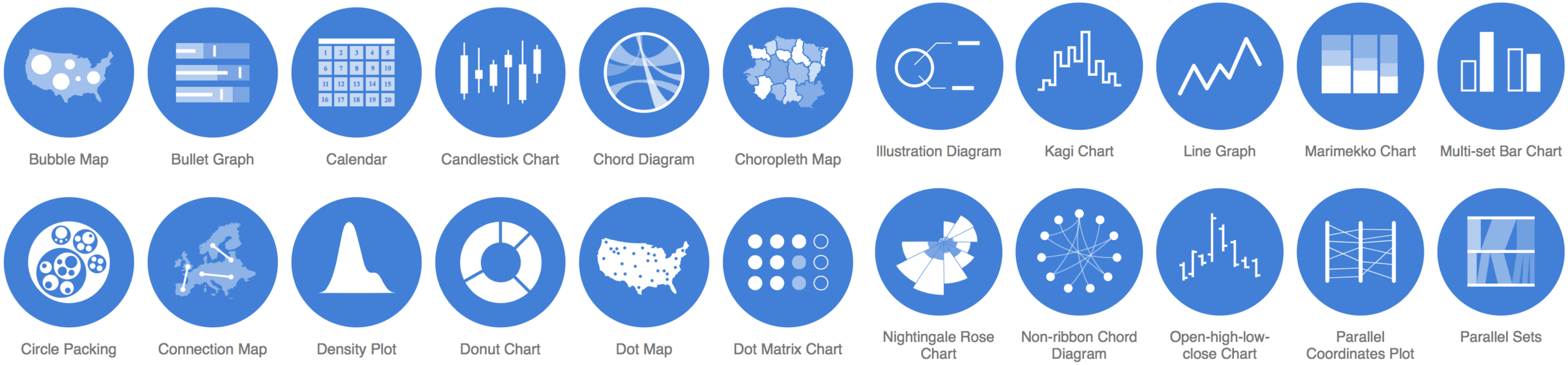


https://www.theguardian.com/cities/gallery/2016/jun/08/how-far-distance-workers-commute-uk-cities-mapped

Intro & Review

ENVS225 Exploring the Social World

Gabriele Filomena
gfilo@liverpool.ac.uk

Data Preparation

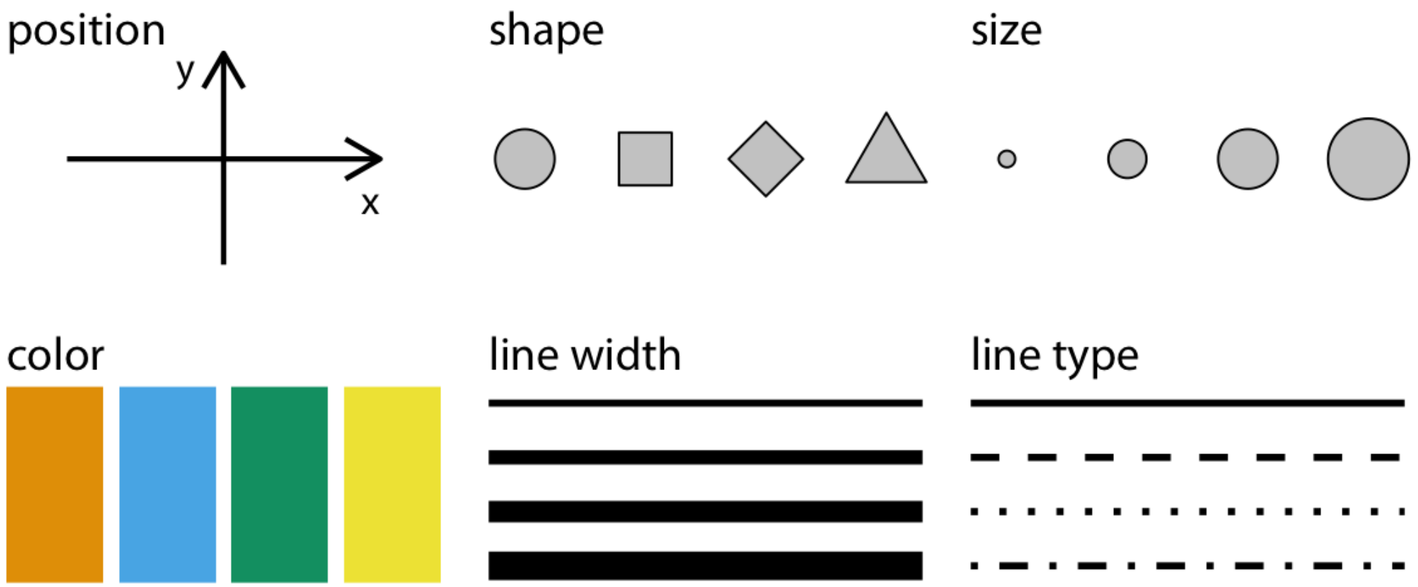
Make sure your data is “tidy”:

- Observations are rows.
- Variables are columns.
- Every cell is a single value.
- First row (and only first row) consists of variable names.

Plots elements

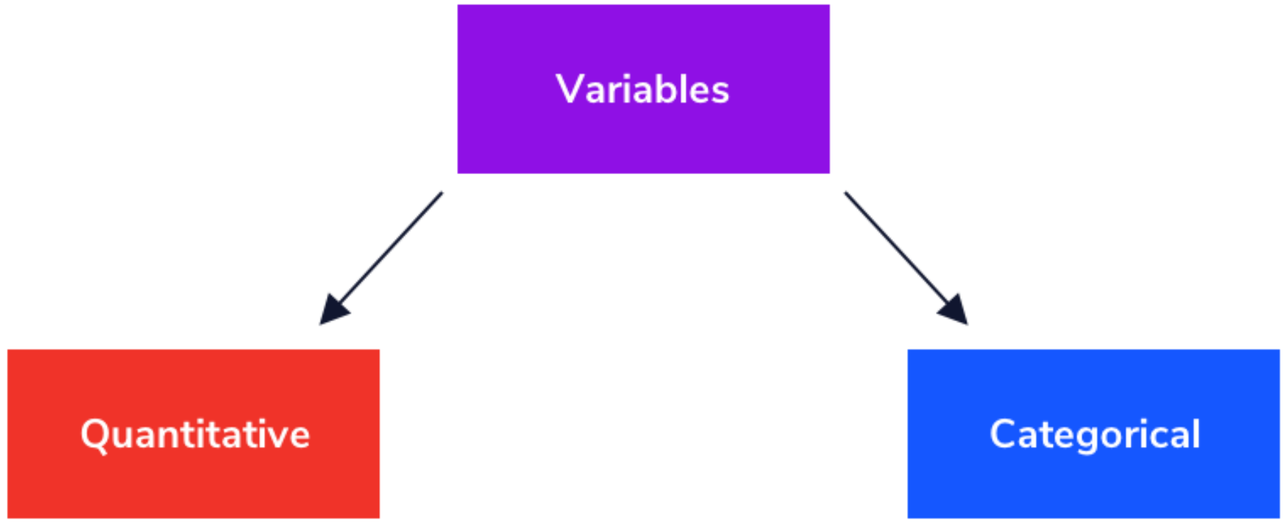
6 elements of a plot that can be assigned to variables:

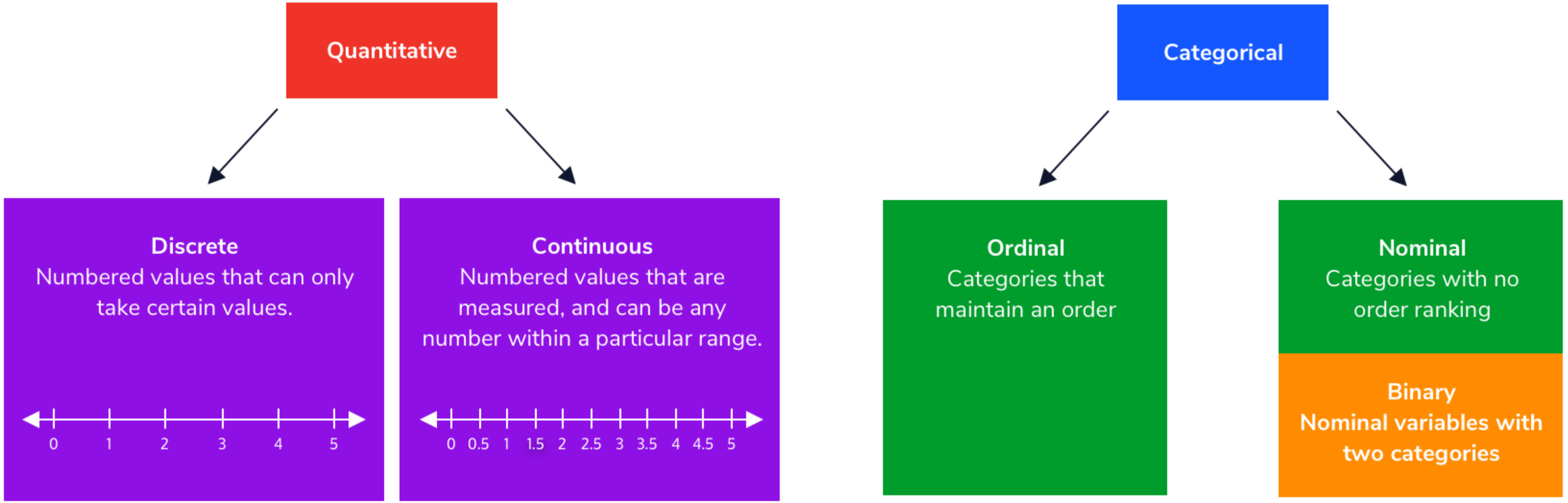
- x-axis (horizontal axis)
- y-axis (vertical axis)
- Facets (types)
- colour/fill
- size
- shape



Variable Types

- Categorical: *colour*, *species*, *sex*.
- Quantitative: *height*, *GDP*, *score*




Quantitative
Discrete
Numbered values that can only take certain values.
0 1 2 3 4 5
Continuous
Numbered values that are measured, and can be any number within a particular range.
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5
Categorical
Ordinal
Categories that maintain an order
Nominal
Categories with no order ranking
Binary
Nominal variables with two categories

Visualising and Describing Variables

What is Data Visualisation?

We take data values and convert them in a systematic and logical way into the visual elements that make up the final graphic.

“A data visualization first and foremost has to accurately convey the data. It must not mislead or distort. At the same time, a data visualization should be aesthetically pleasing”

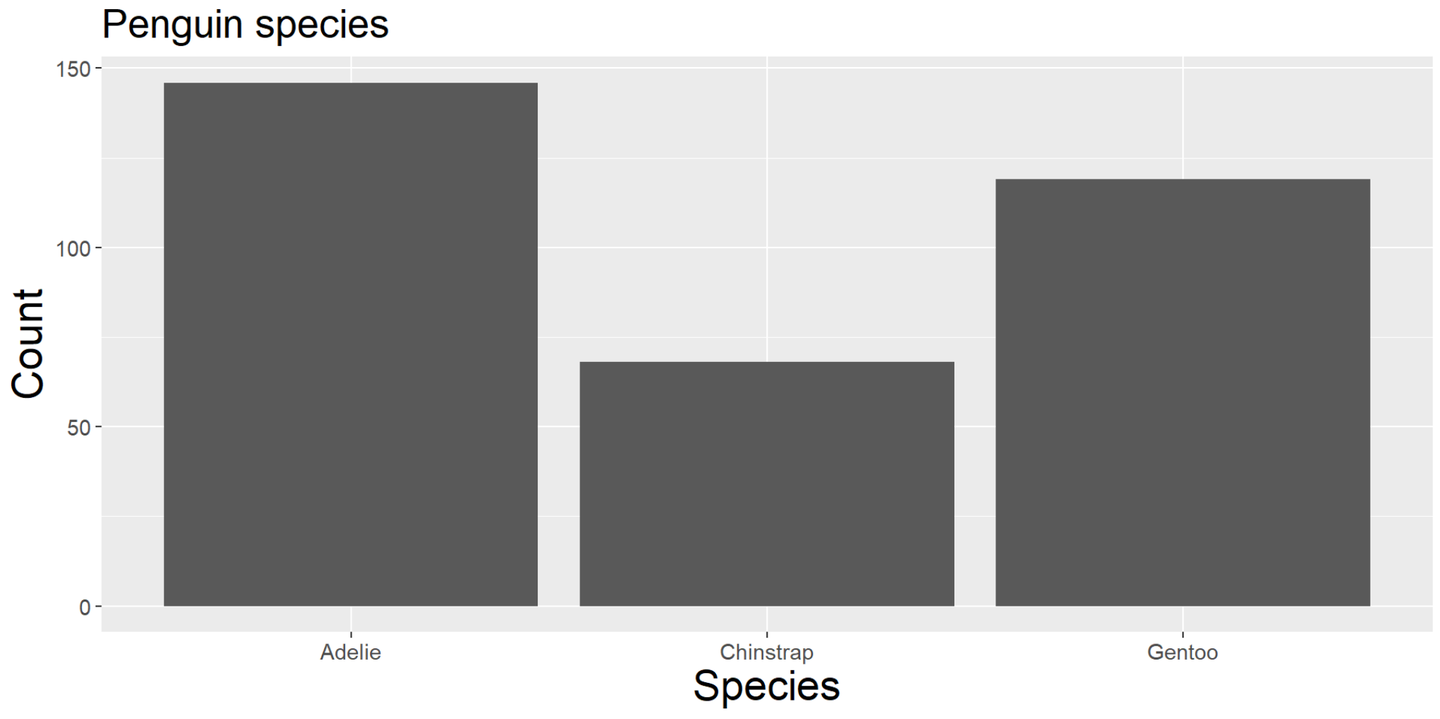
Key Principle

Edward Tufte (1983) :
“What is to be sought in designs for the display of information is the clear portrayal of complexity”

→ More is less

1 Numerical Variable:

Histogram

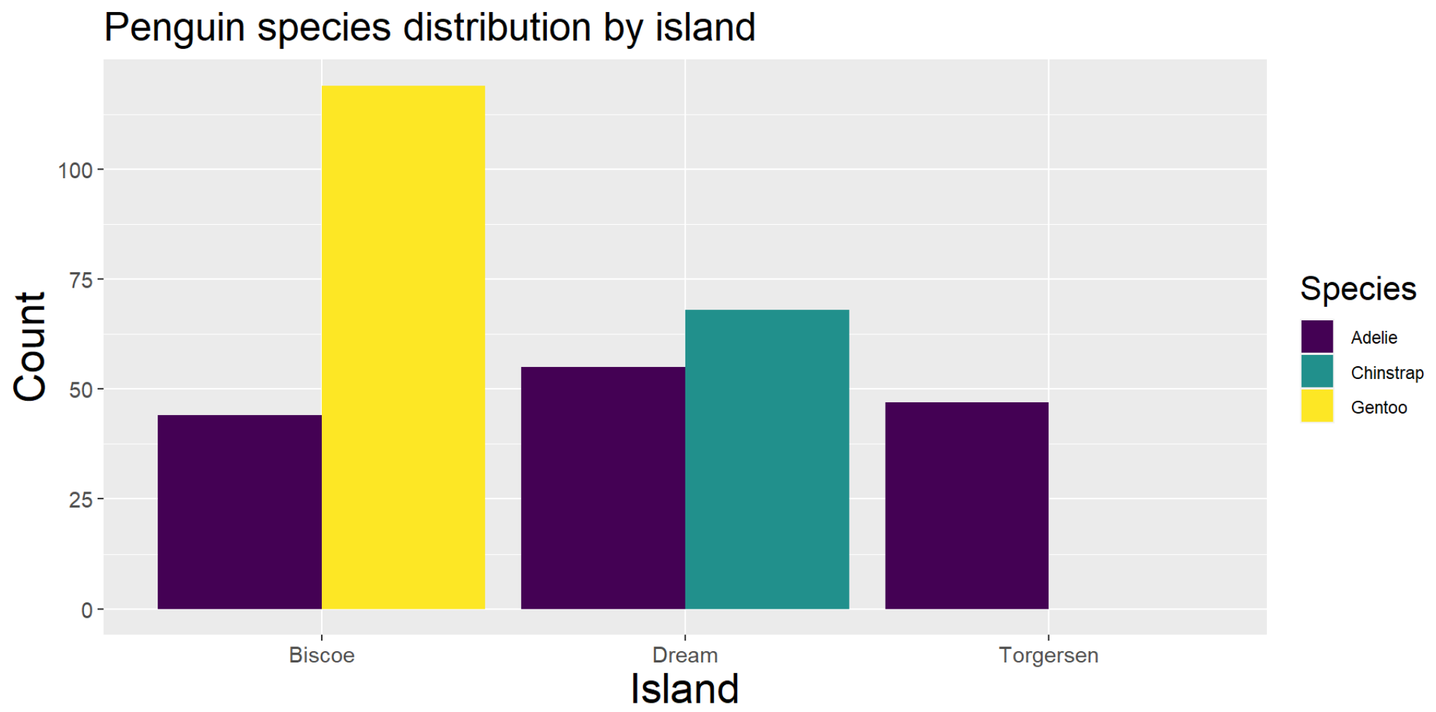


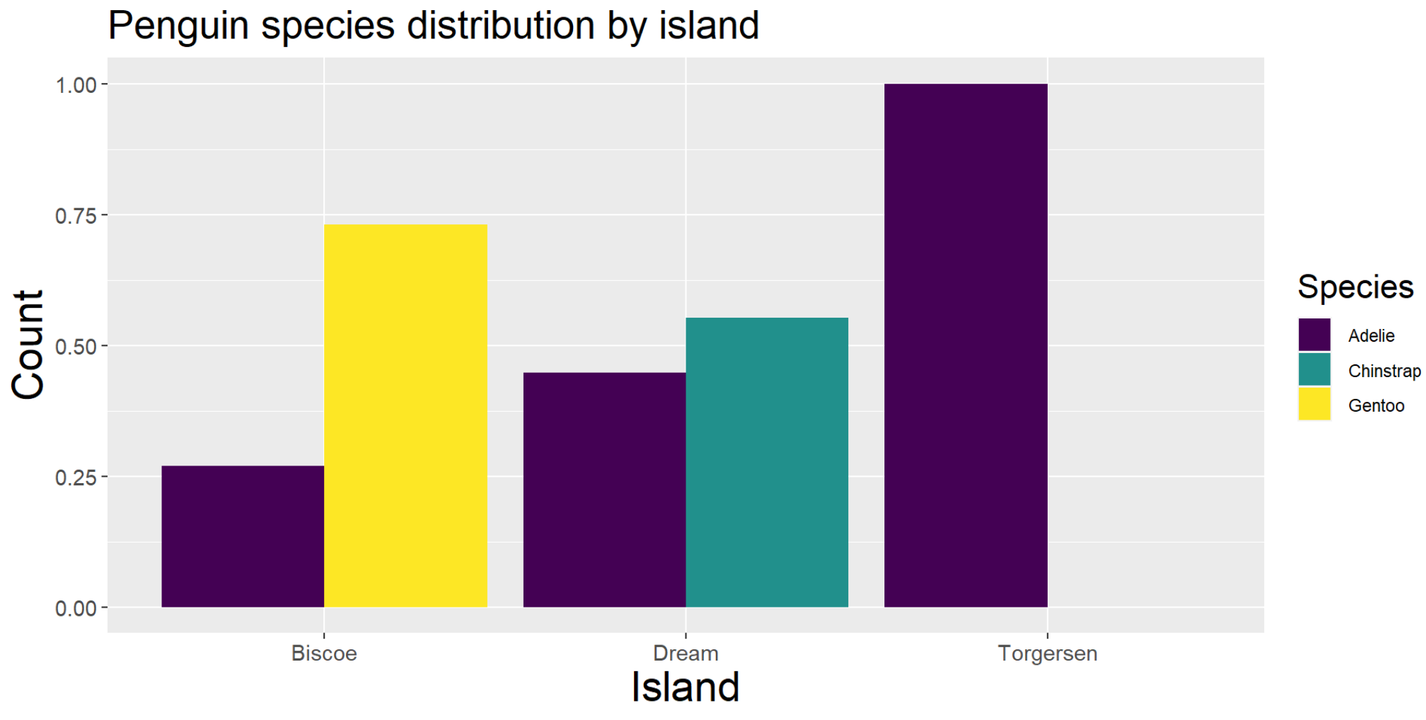
1 Categorical Variable:

Bar Chart

2 Categorical Variables:

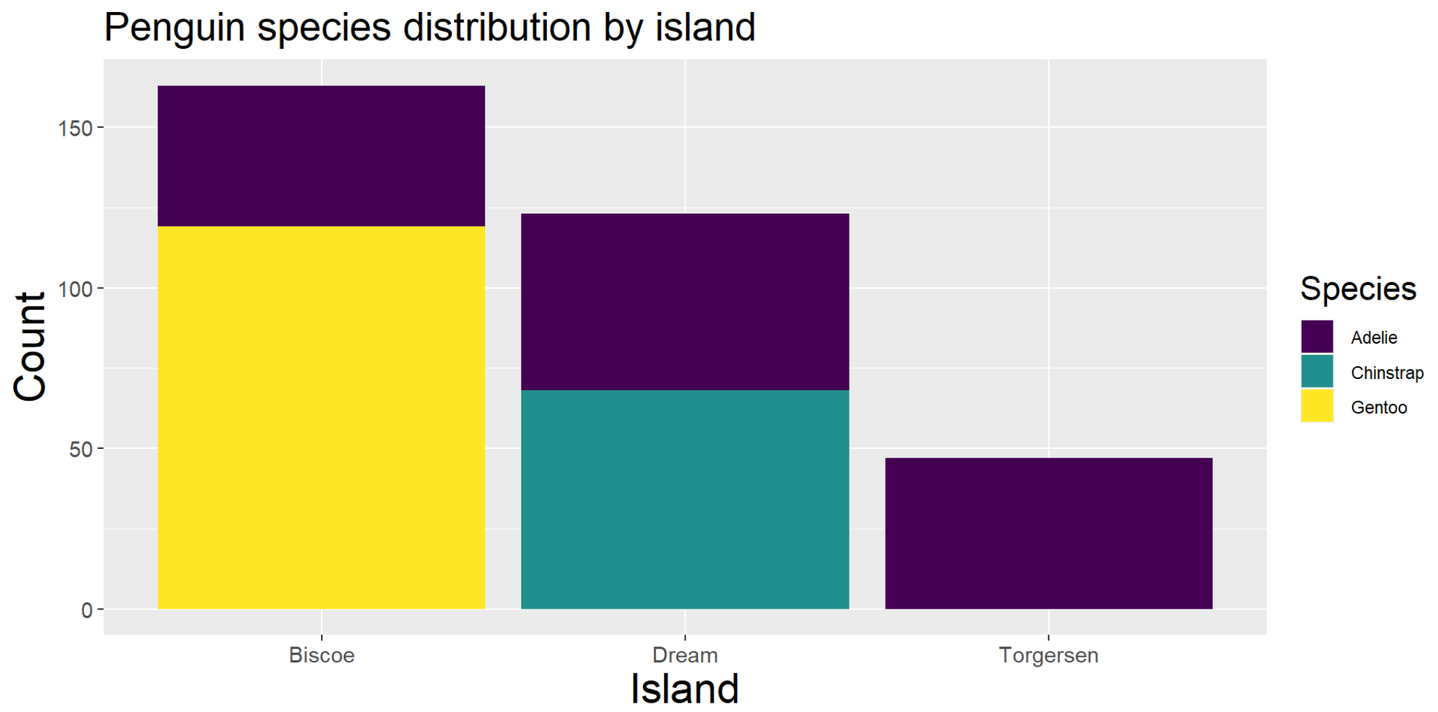
Side-by-side bar chart

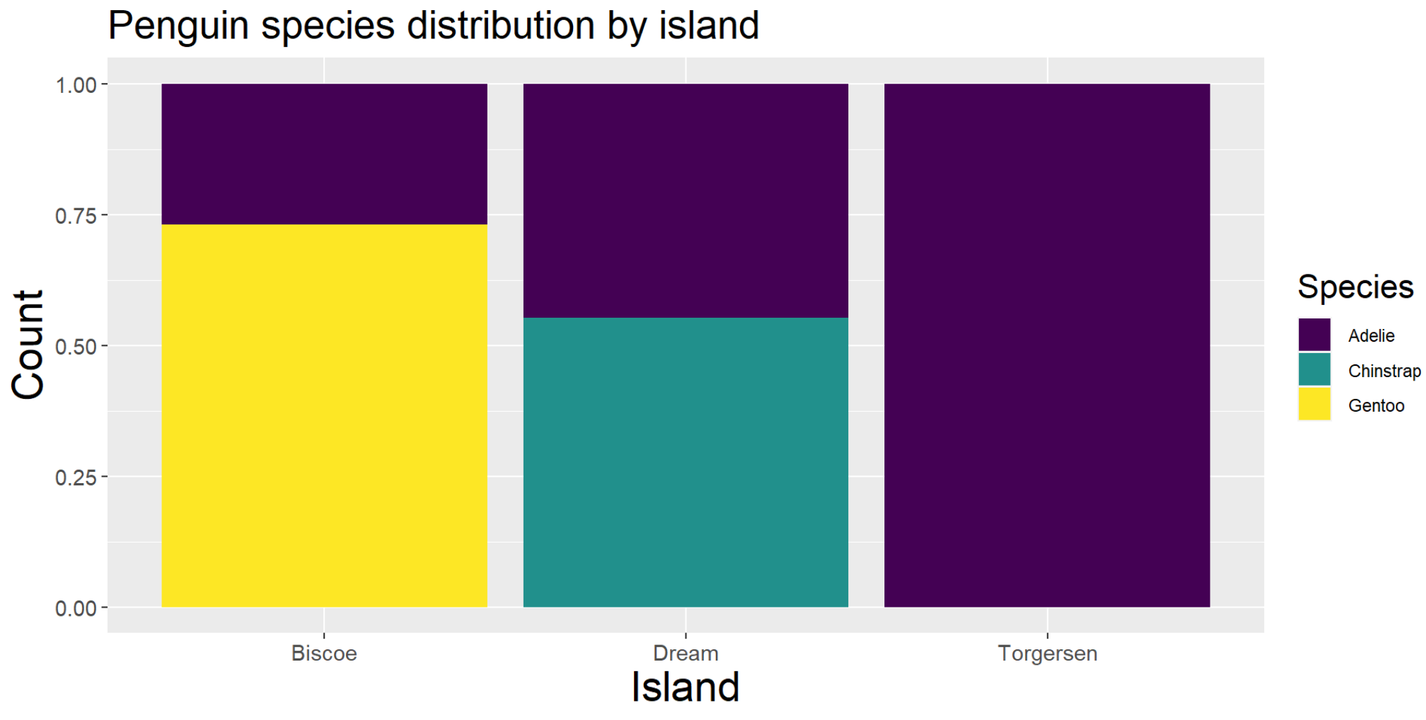




2 Categorical Variables:

Stacked bar chart

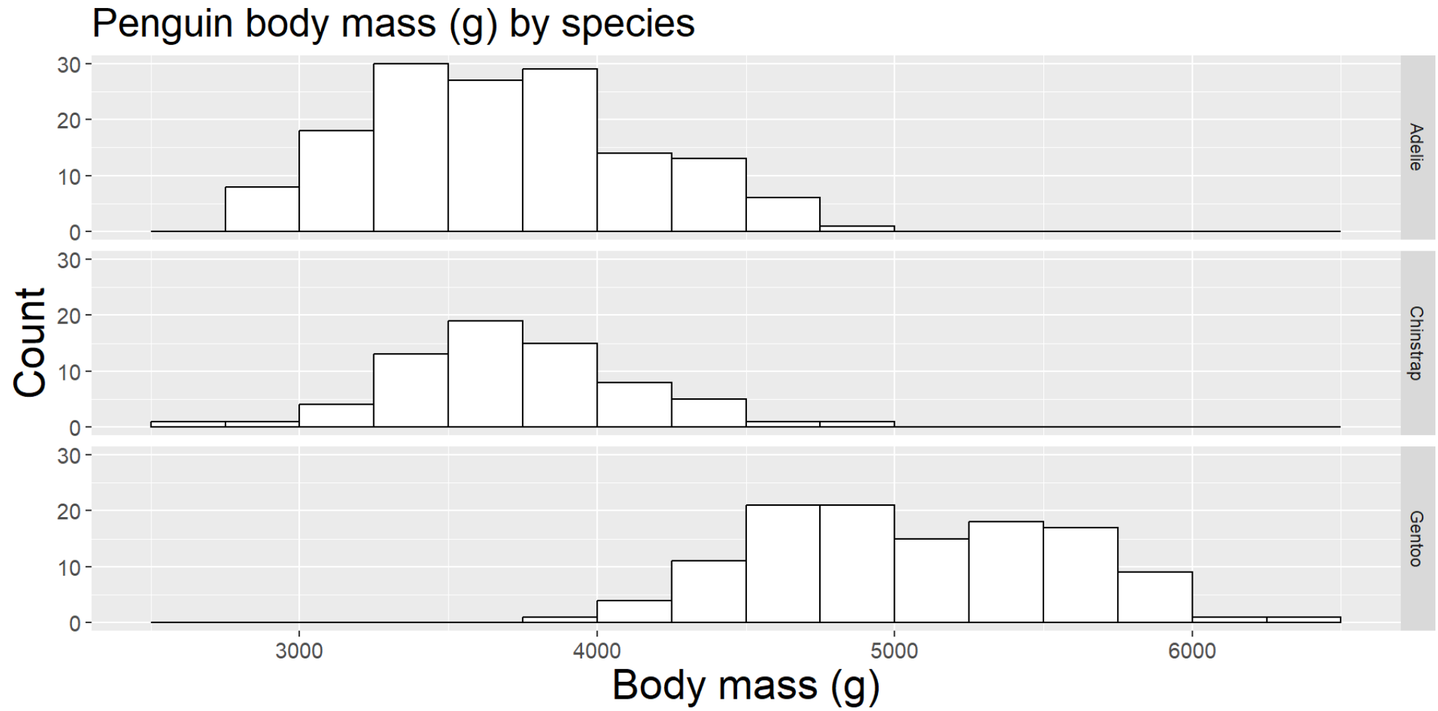


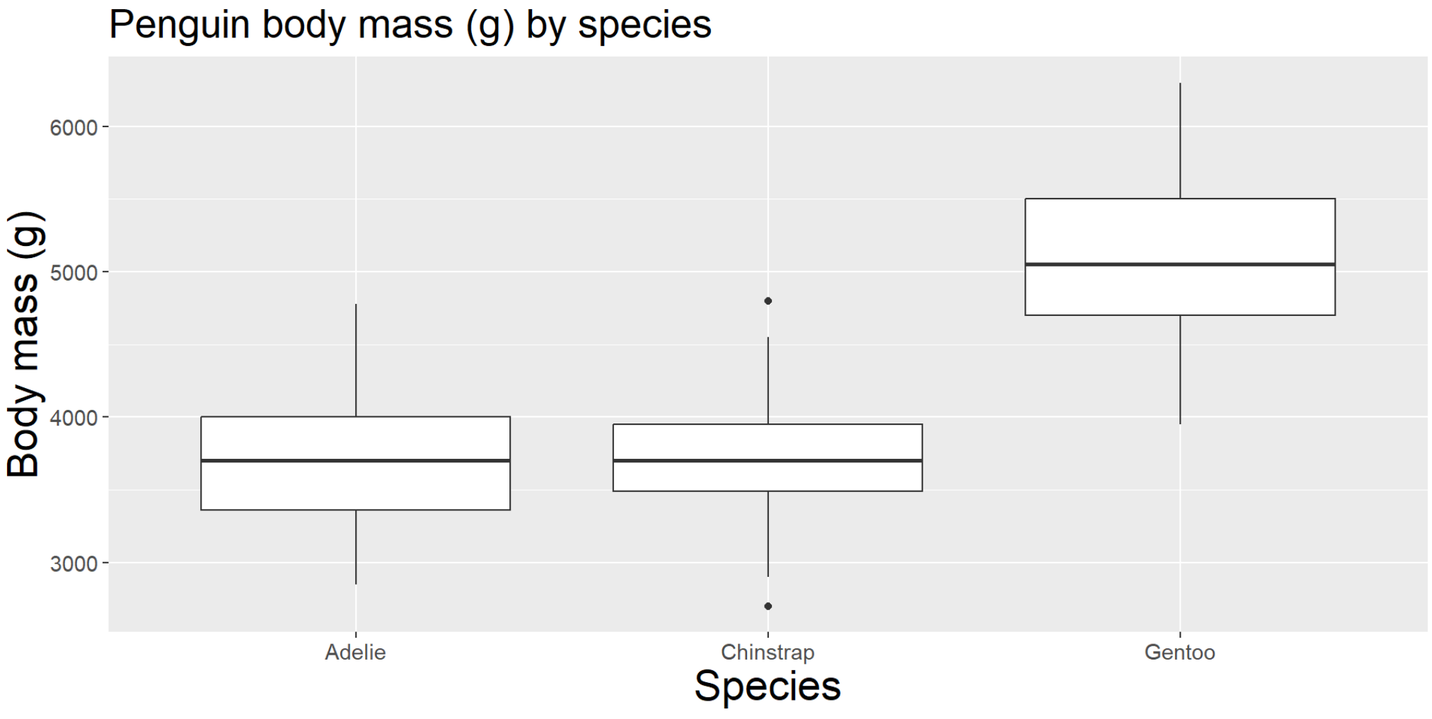


1 Categorical Variable
1 Numerical Variable

Side-by-side boxplot (bottom)

Stacked histogram (right)

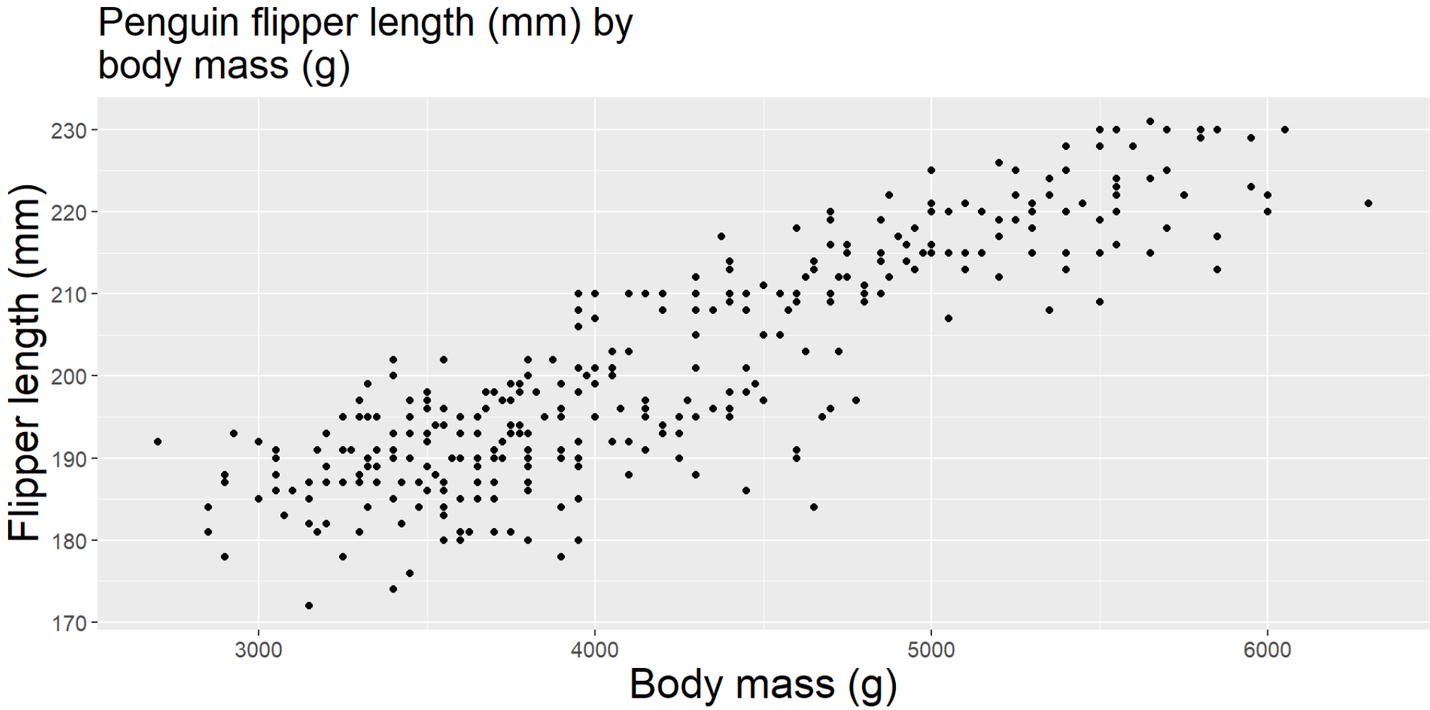


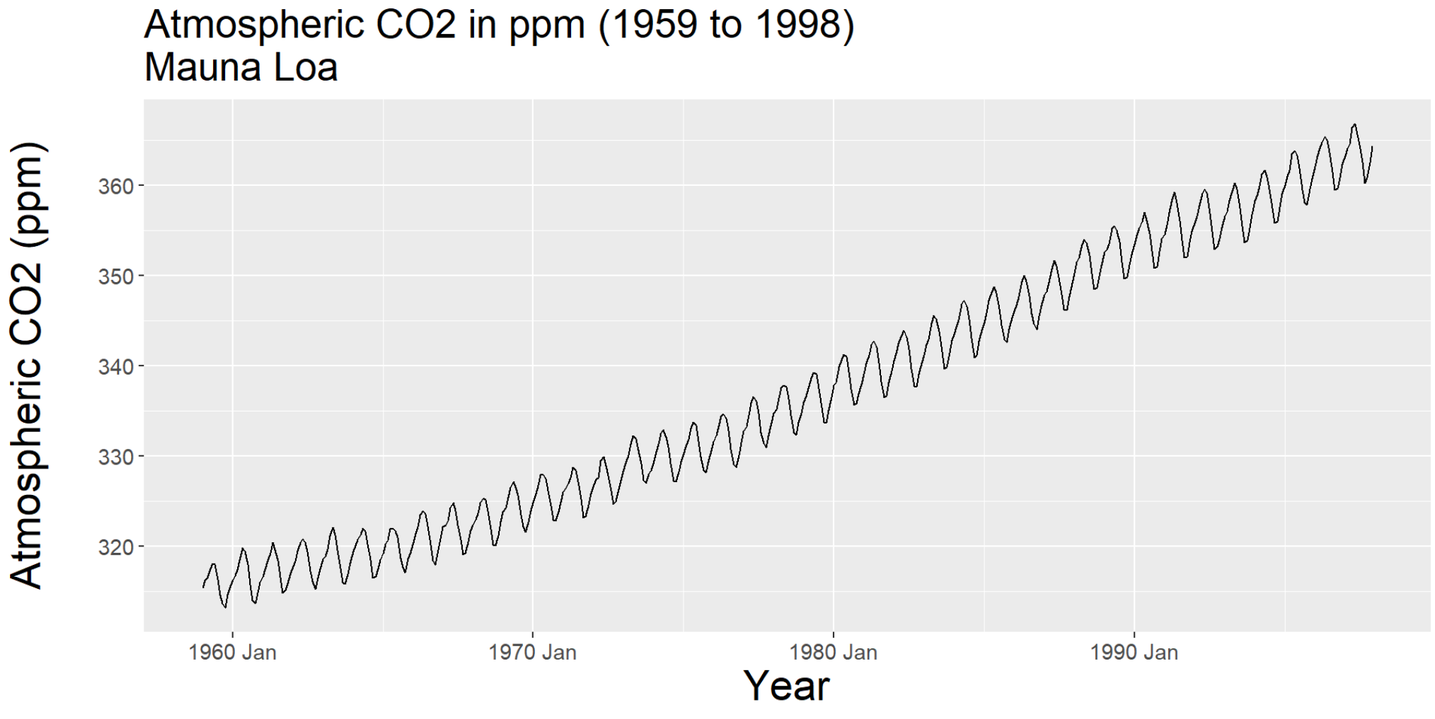


2 Numerical Variables

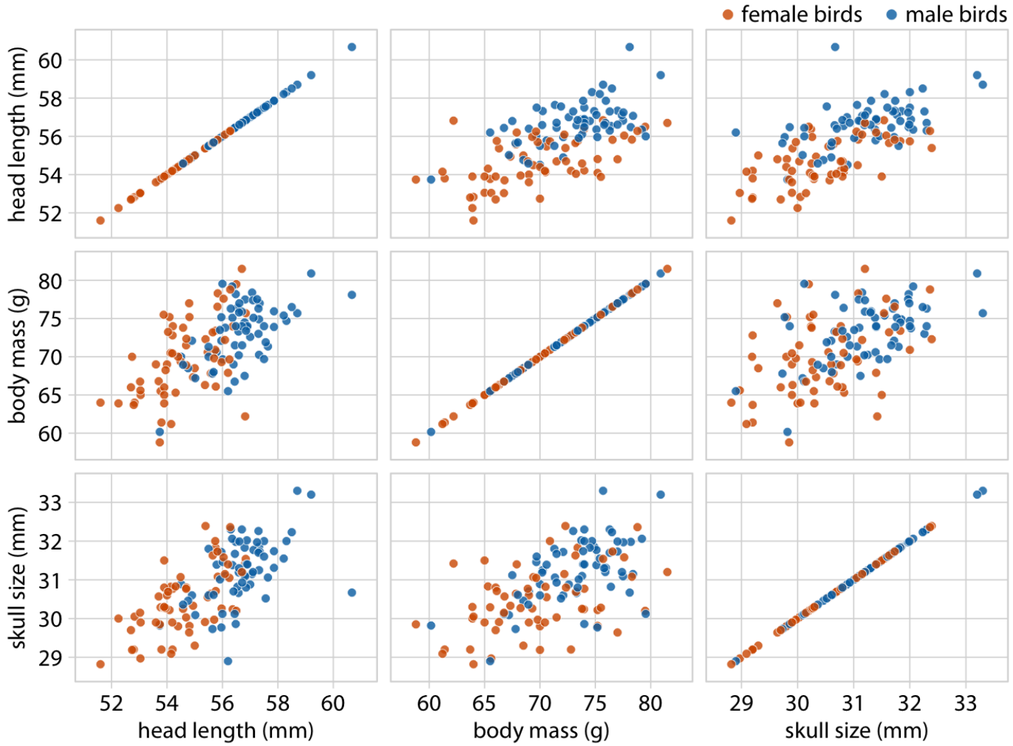
Scatterplot (right)

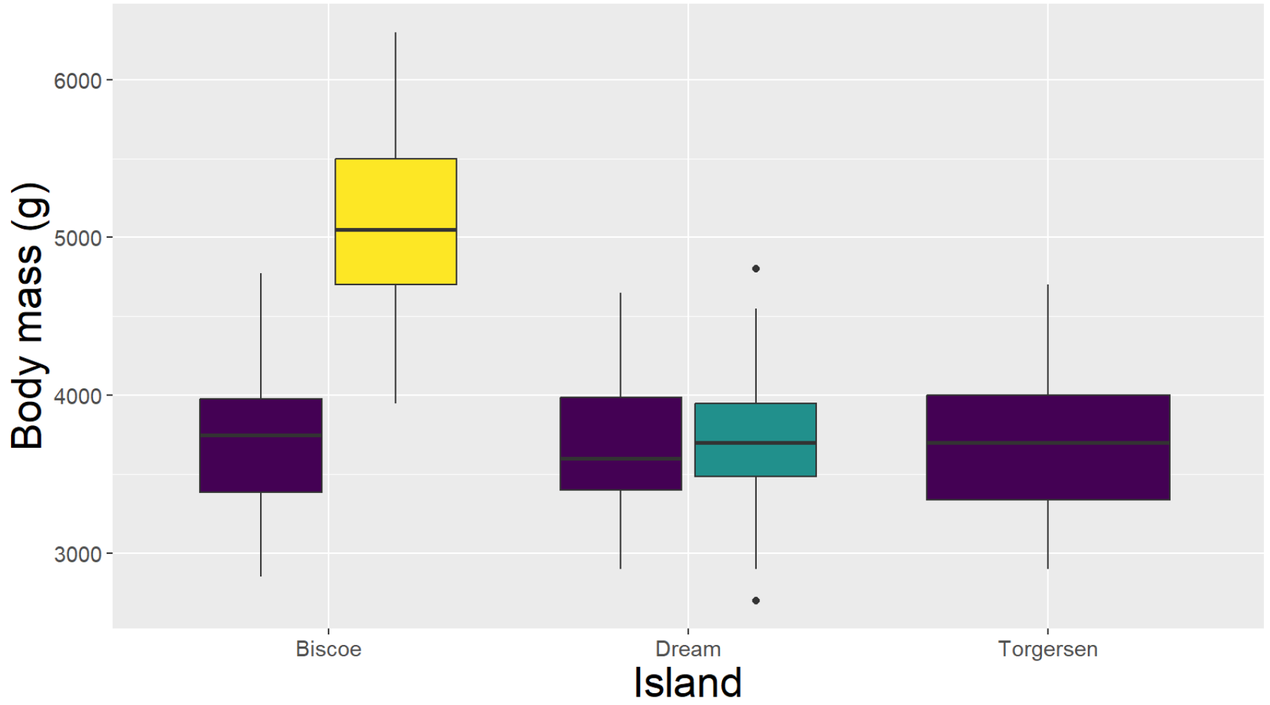
Lineplot (bottom)

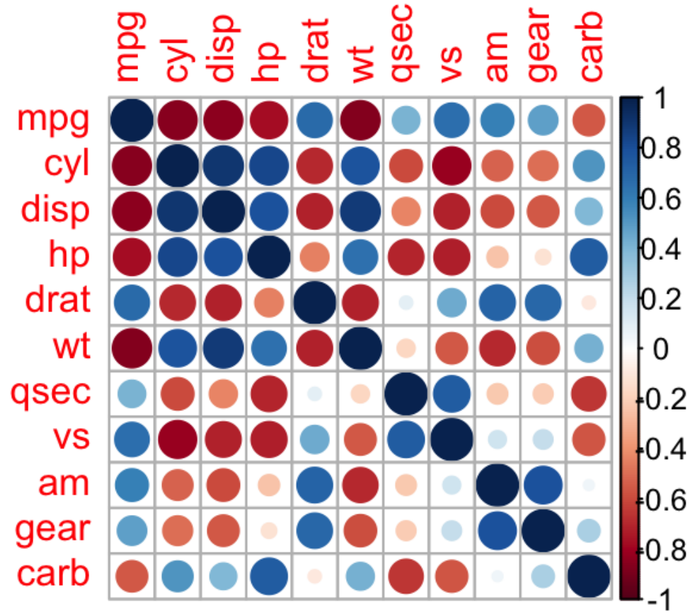




3+ Variables

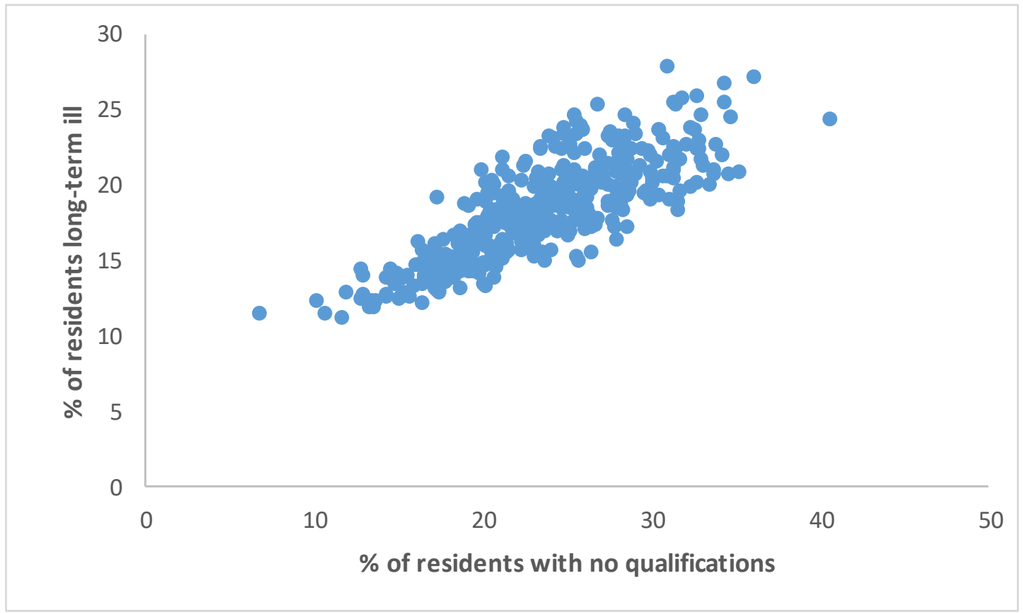

female birds
male birds
head length (mm)
body mass (g)
skull size (mm)


Body mass (g)
Island
Biscoe
Dream
Torgersen
Species
Adelie
Chinstrap
Gentoo

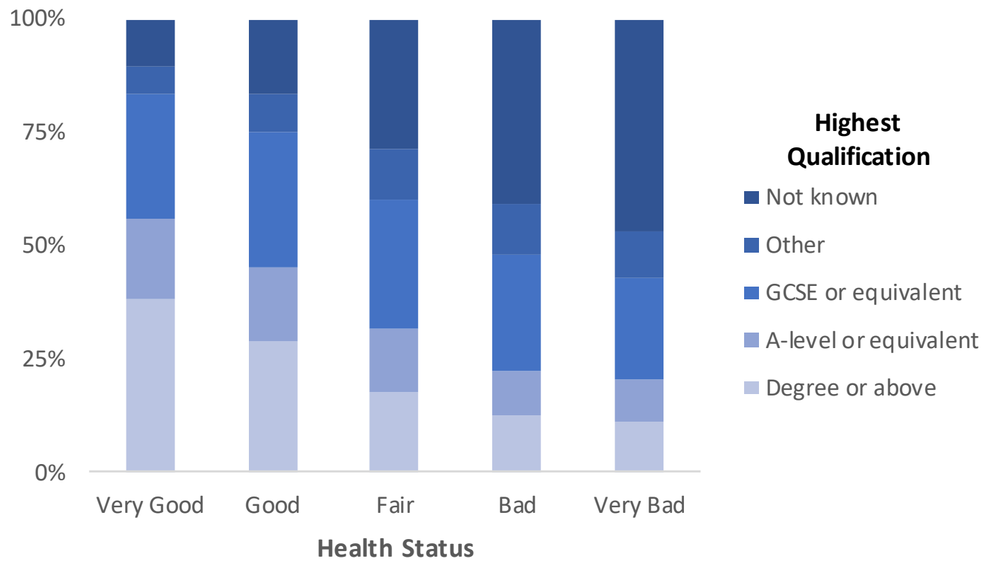

mpg
cyl
disp
hp
drat
wt
qsec
vs
am
gear
carb

Visualising Relationships

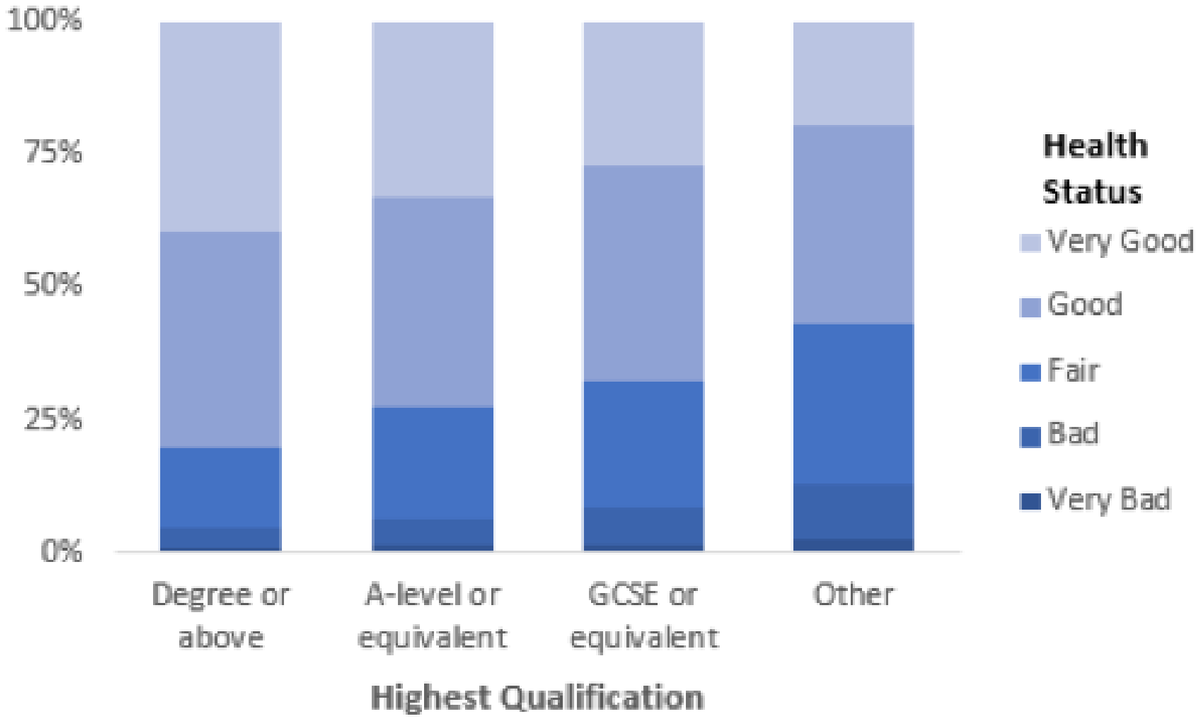
For continuous data:
Scatter-plots



For categorical data:
Stacked percentage bar-charts



Key Elements



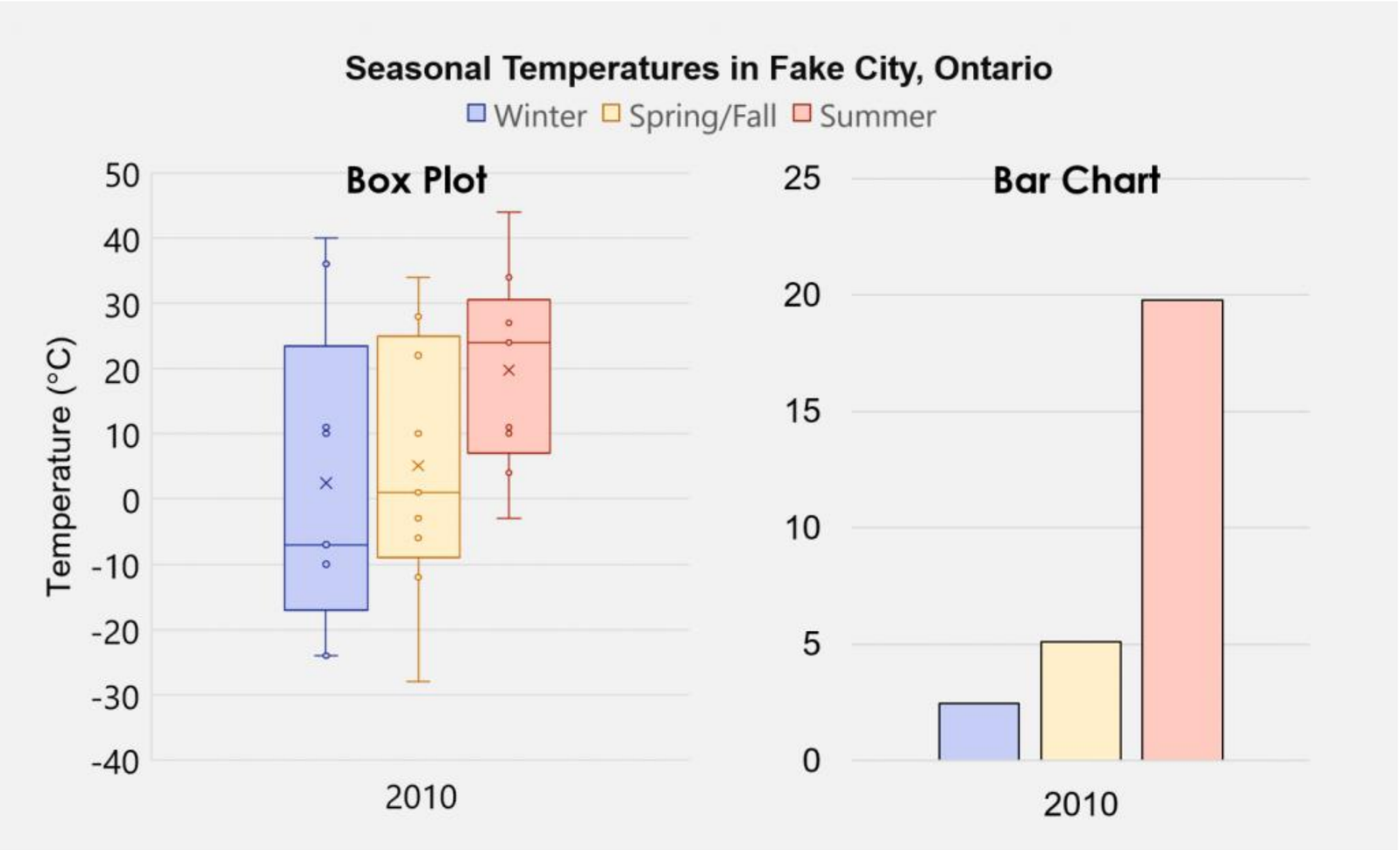
- Title (Descriptive)
- Number
- Data source (if necessary)
- Axis labels & titles
- Legend
- Clear colours

Figure 1 Health and Highest Qualification

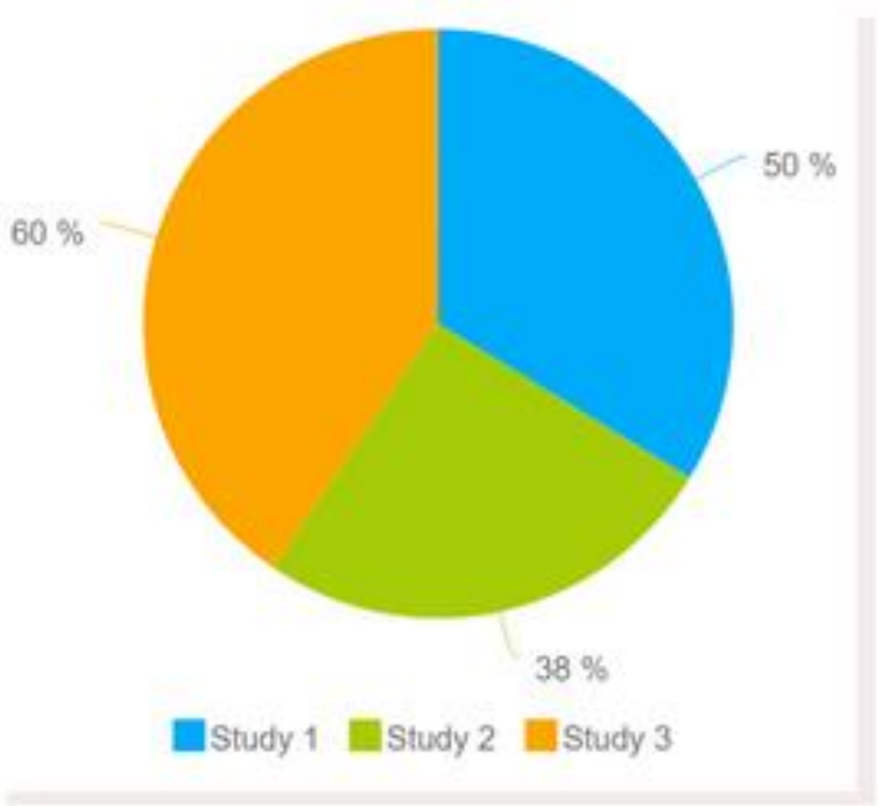
Source: Author's calculations based upon the *Family Resources Survey 2016-17* (DWP/NCSR/ONS, 2018).

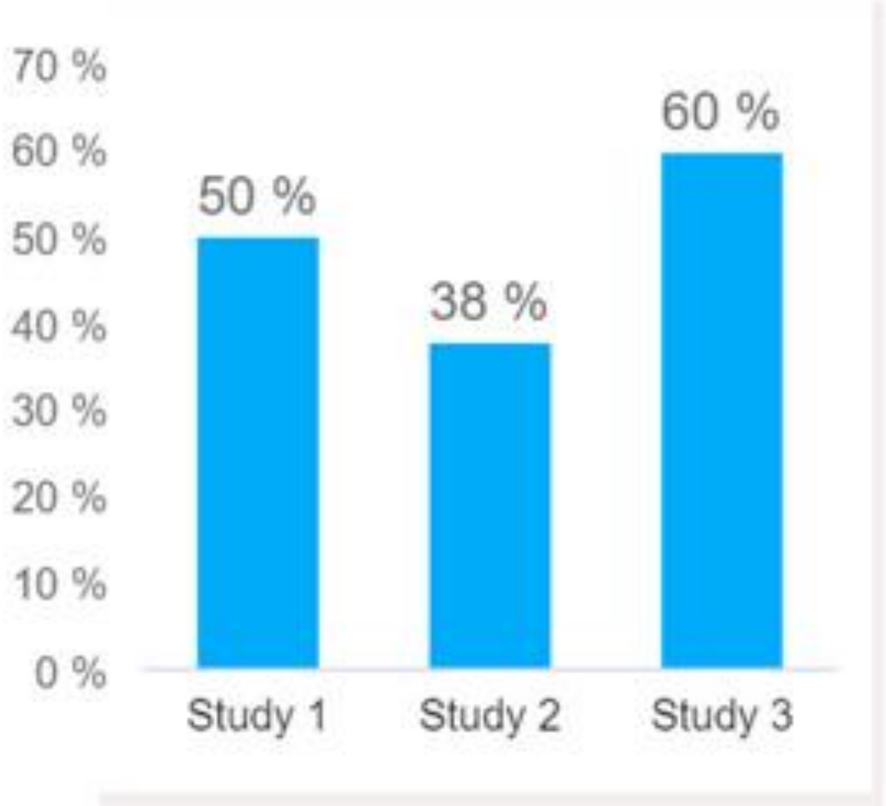
Bad and Good

Representing Distributions

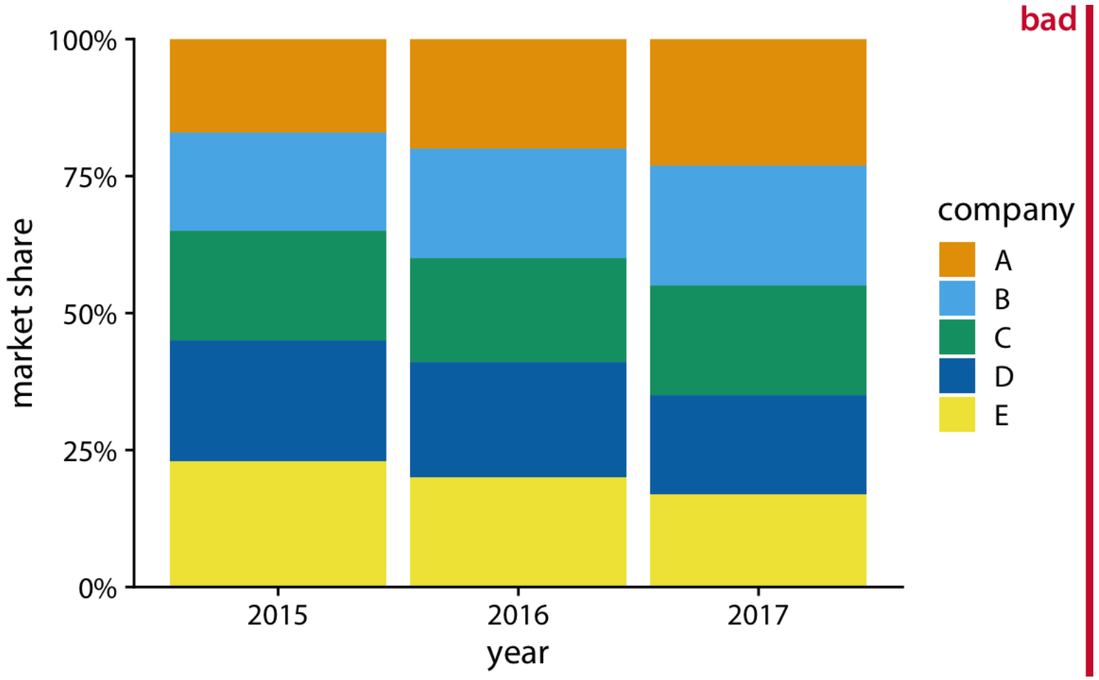


Percentages/ Proportions

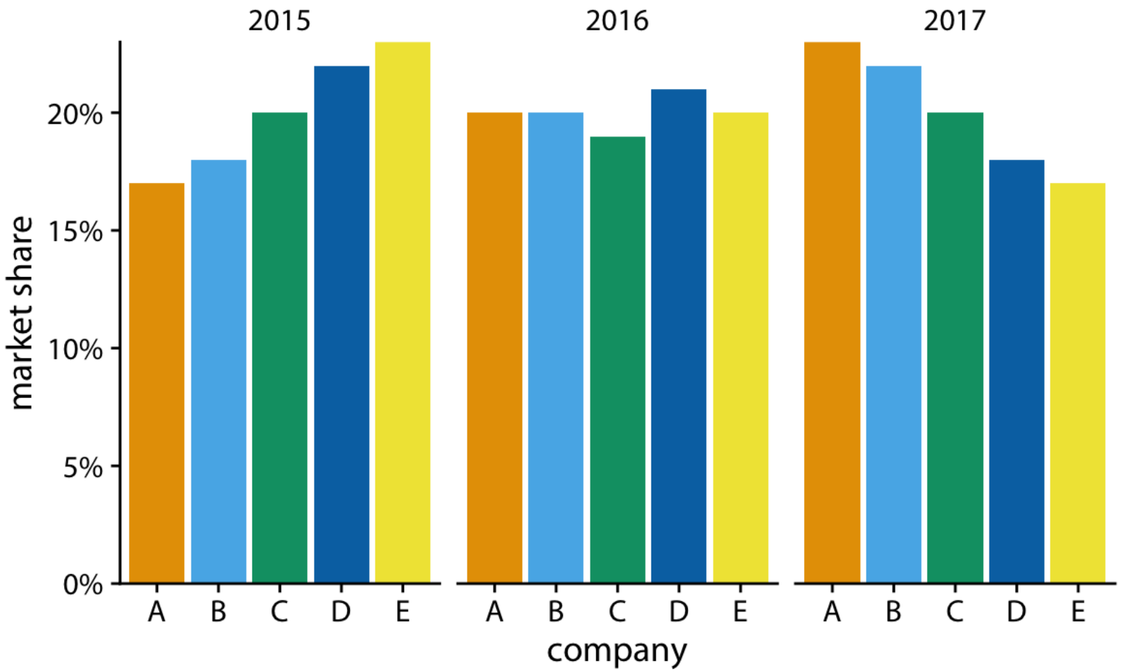




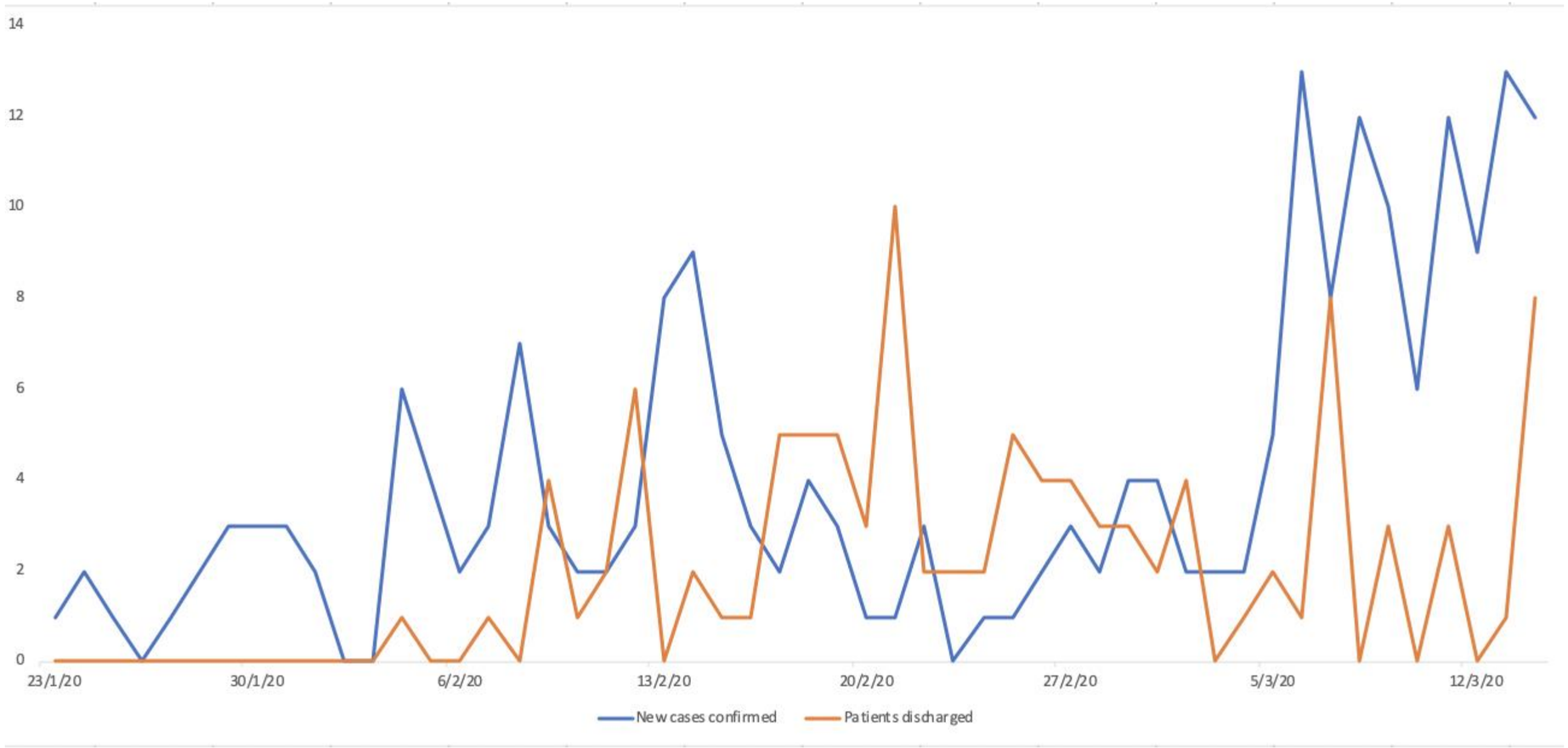
Percentages/ Proportions



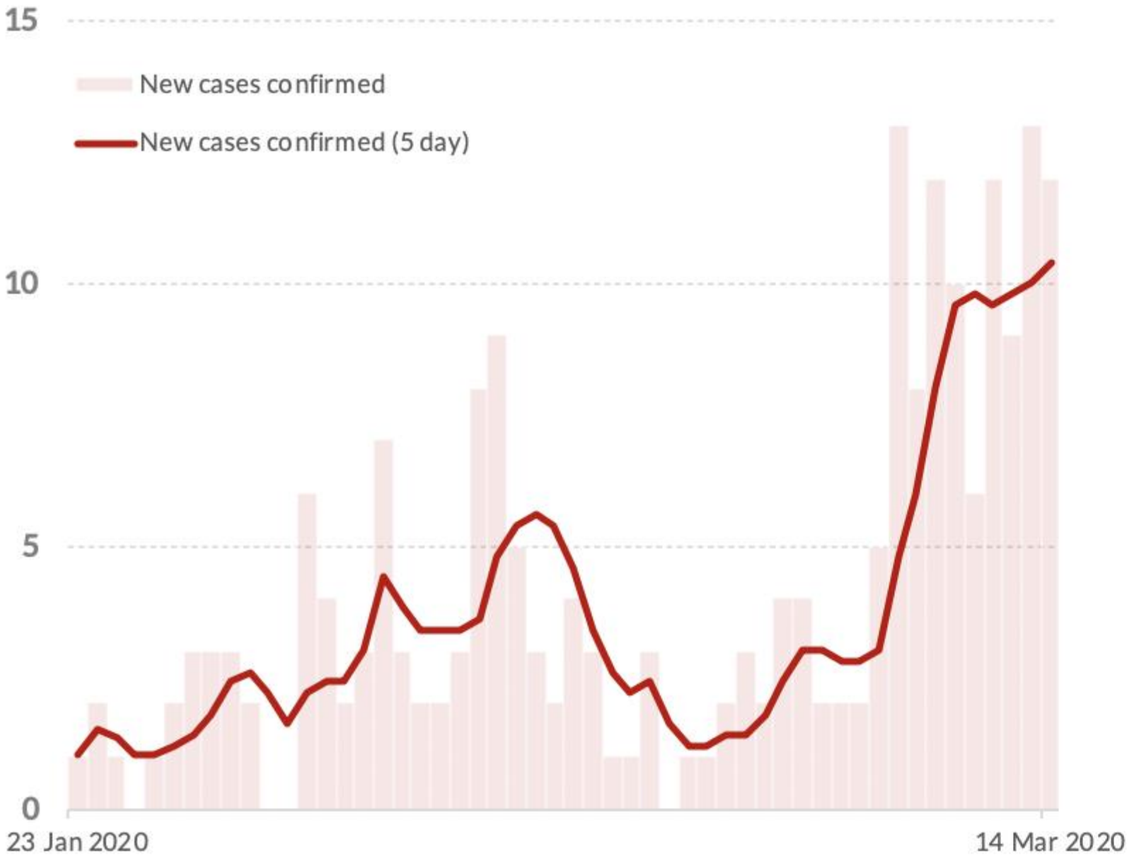
bad

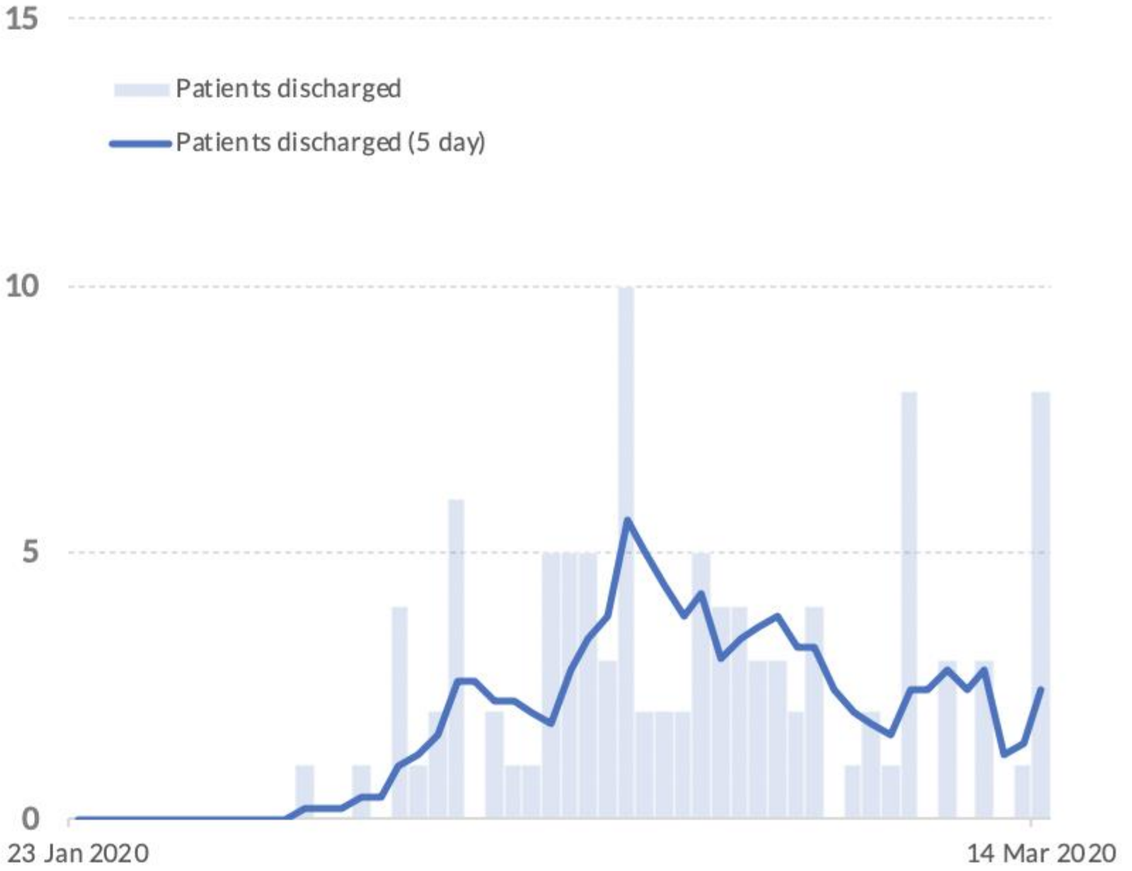


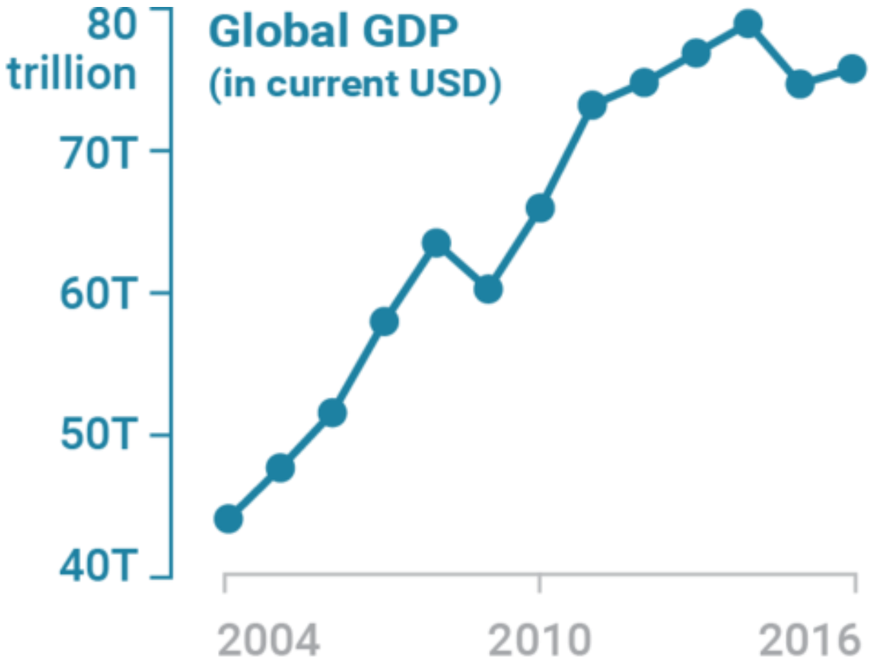
Lineplots?

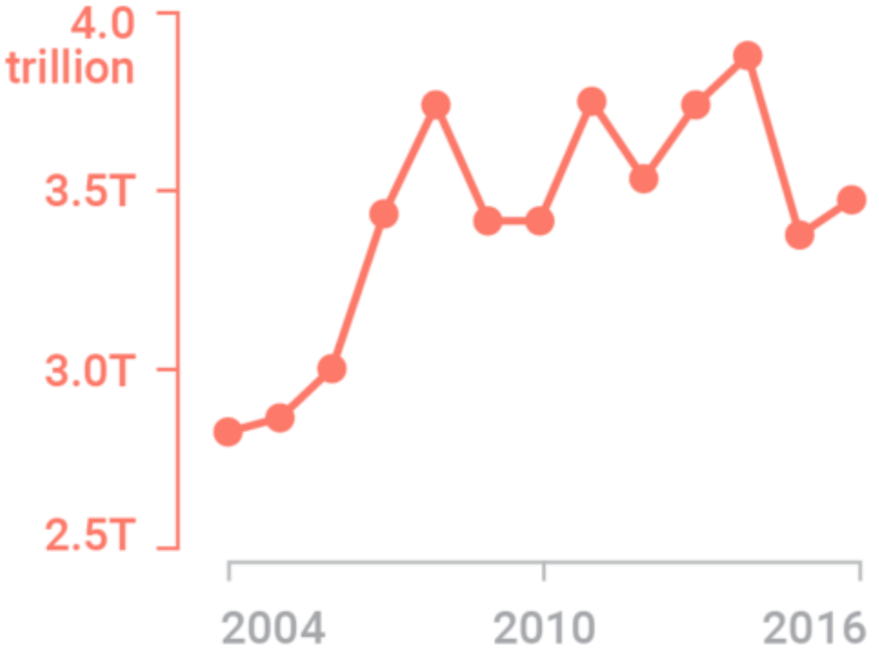

14
12
10
8
6
4
2
0
23/1/20
30/1/20
6/2/20
13/2/20
20/2/20
27/2/20
5/3/20
12/3/20
New cases confirmed
Patients discharged

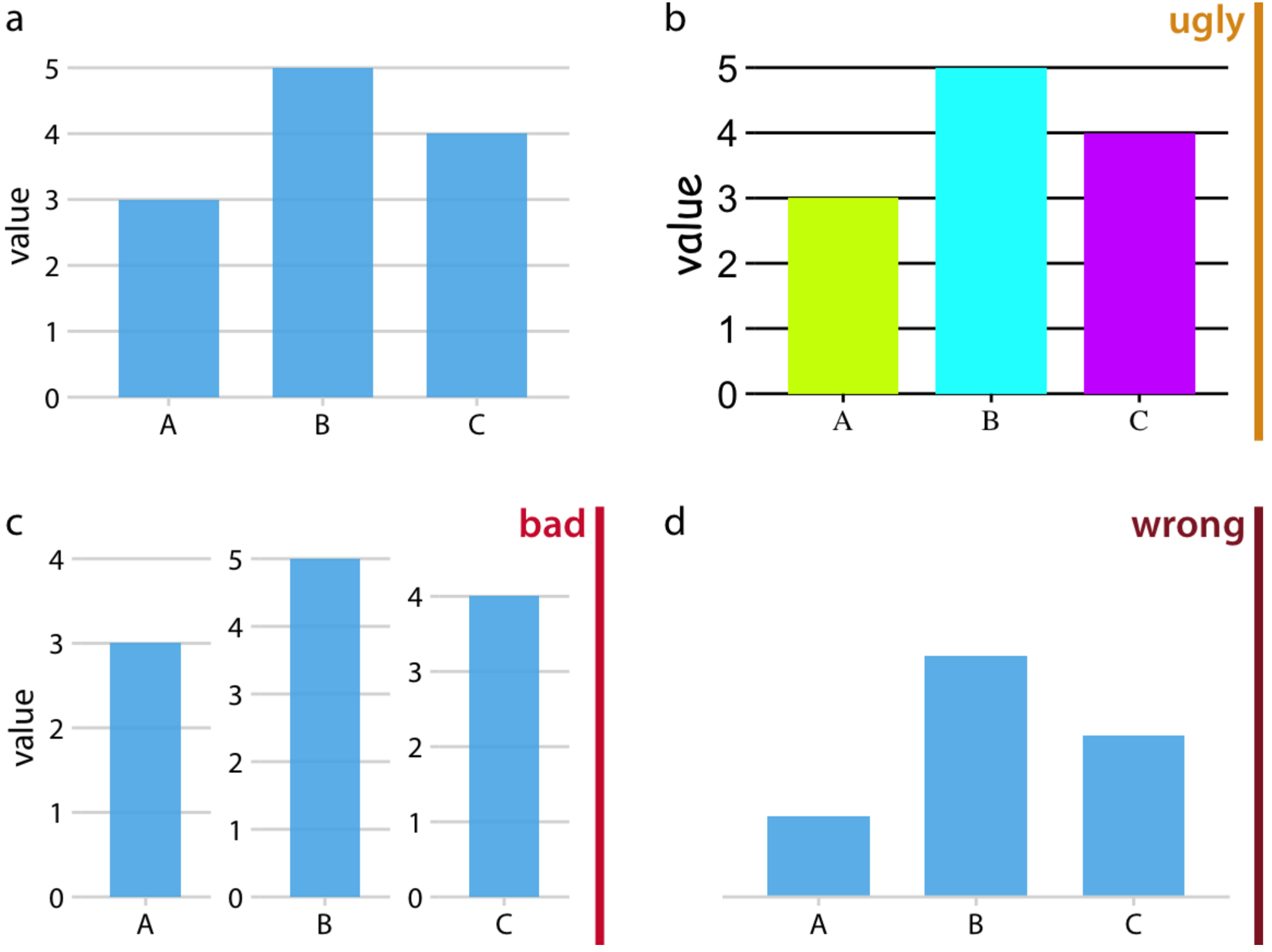
COVID-19 Case in Singapore: New Cases vs Newly Discharged






80
trillion
70T
60T
50T
40T
Global GDP
(in current USD)
2004
2010
2016


4.0
trillion
3.5T
3.0T
2.5T
2004
2010
2016


a
value
5
4
3
2
1
0
A
B
C
b
ugly
value
5
4
3
2
1
0
A
B
C
c
bad
4
3
2
1
0
5
4
3
2
1
0
4
3
2
1
0
A
B
C
value
d
wrong
A
B
C

Tables

a

ugly

Rank	Title	Amount
1	*Star Wars: The Last Jedi*	$71,565,498
2	*Jumanji: Welcome to the Jungle*	$36,169,328
3	*Pitch Perfect 3*	$19,928,525
4	*The Greatest Showman*	$8,805,843
5	*Ferdinand*	$7,316,746

b

ugly

Rank	Title	Amount
1	*Star Wars: The Last Jedi*	$71,565,498
2	*Jumanji: Welcome to the Jungle*	$36,169,328
3	*Pitch Perfect 3*	$19,928,525
4	*The Greatest Showman*	$8,805,843
5	*Ferdinand*	$7,316,746

c

Rank	Title	Amount
1	*Star Wars: The Last Jedi*	$71,565,498
2	*Jumanji: Welcome to the Jungle*	$36,169,328
3	*Pitch Perfect 3*	$19,928,525
4	*The Greatest Showman*	$8,805,843
5	*Ferdinand*	$7,316,746

d

Rank	Title	Amount
1	*Star Wars: The Last Jedi*	$71,565,498
2	*Jumanji: Welcome to the Jungle*	$36,169,328
3	*Pitch Perfect 3*	$19,928,525
4	*The Greatest Showman*	$8,805,843
5	*Ferdinand*	$7,316,746

DOs and DON'Ts

- Avoid overwhelming and poorly designed graphs.
- Keep it simple and clean
- Do not show the same results both with tables and graphs.
- Choose the right chart type.

Sources & Credits

- Graph Types, Sean Raleigh
  Westminster University https://rpubs.com/VectorPosse/graph_types
- Fundamentals of Data Visualization," published by O'Reilly Media, Inc
  https://clauswilke.com/dataviz/index.html