

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

Data visualisation

ENVS225 Exploring the Social World

Gabriele Filomena

gfilo@liverpool.ac.uk

Data Preparation

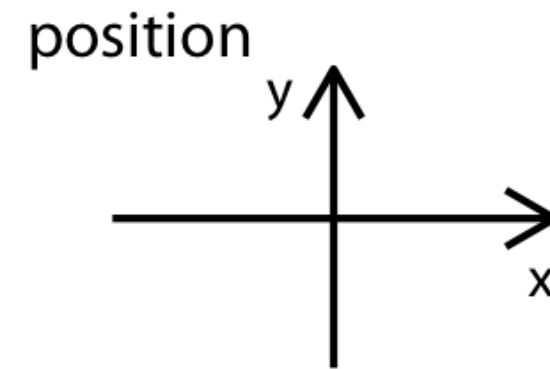
Make sure your data is “tidy”:

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

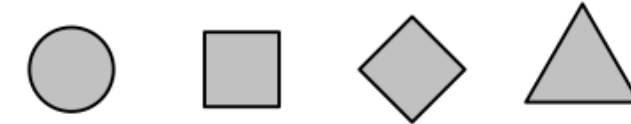
Plots elements

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



shape



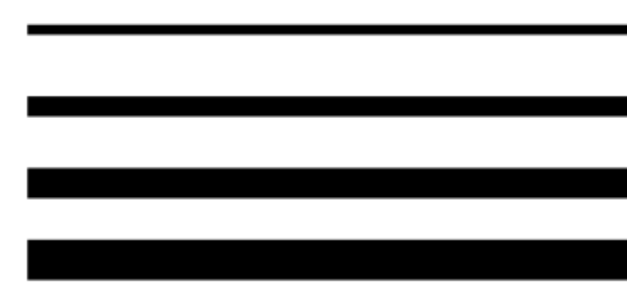
size



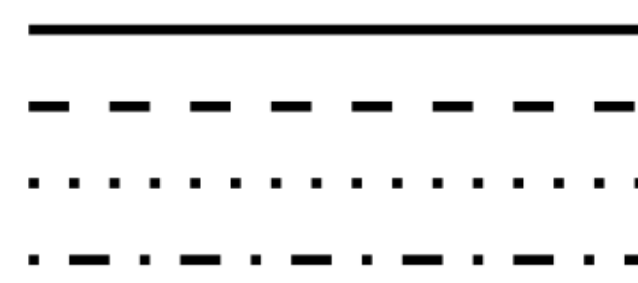
color



line width

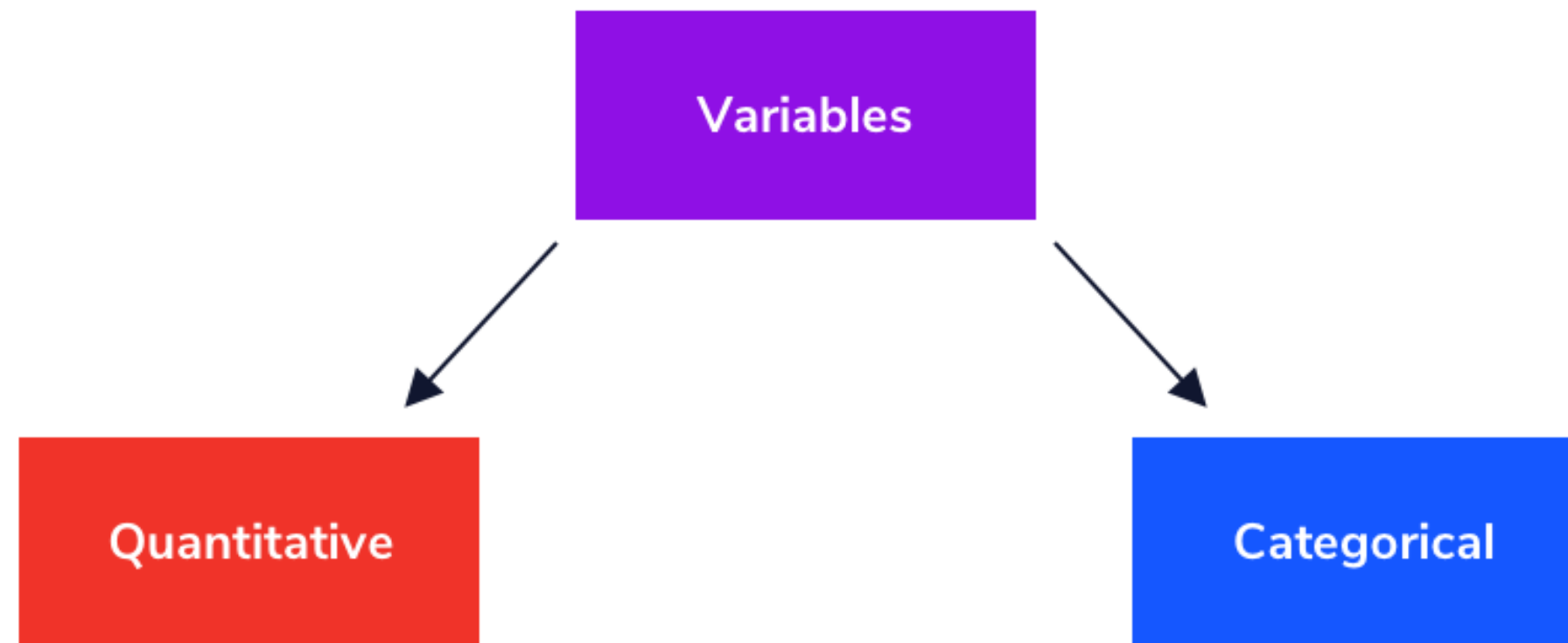


line type



Variable Types

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



Quantitative



Discrete

Numbered values that can only take certain values.



Continuous

Numbered values that are measured, and can be any number within a particular range.



Categorical



Ordinal

Categories that maintain an order

Nominal

Categories with no order ranking

Binary

Nominal variables with two categories

Visualising and Describing Variables

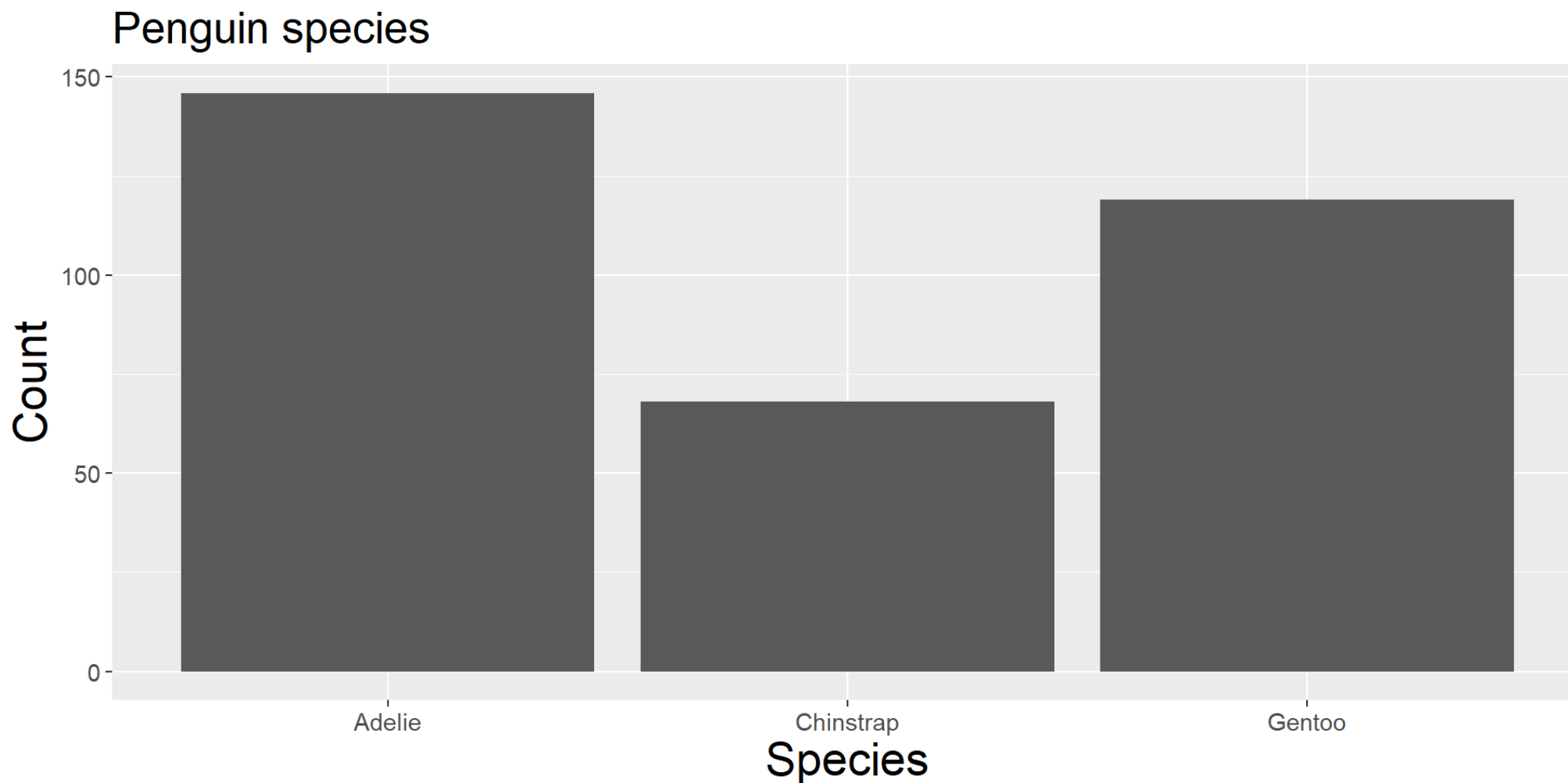
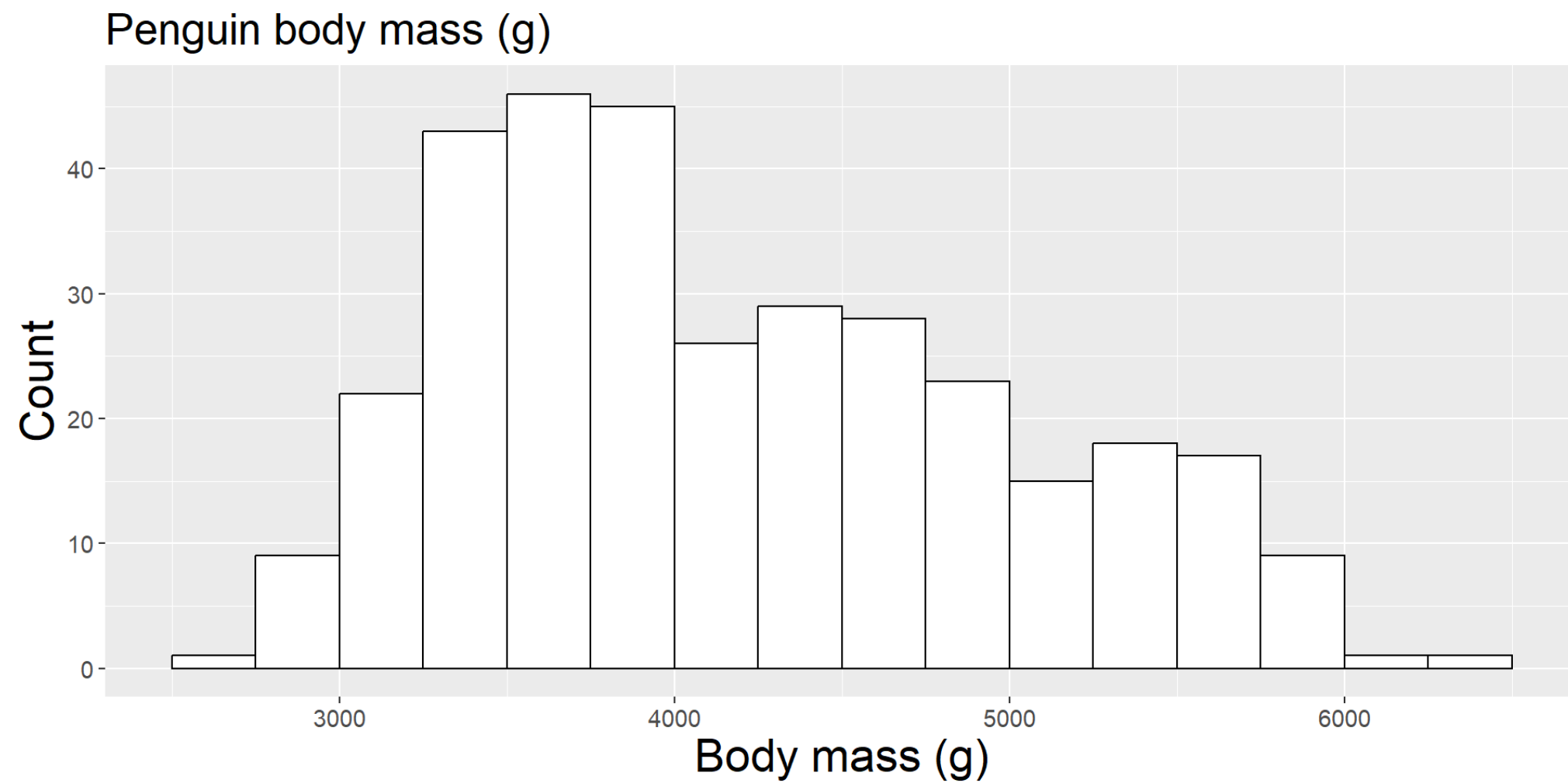
What is Data Visualisation?

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”

→ More is less

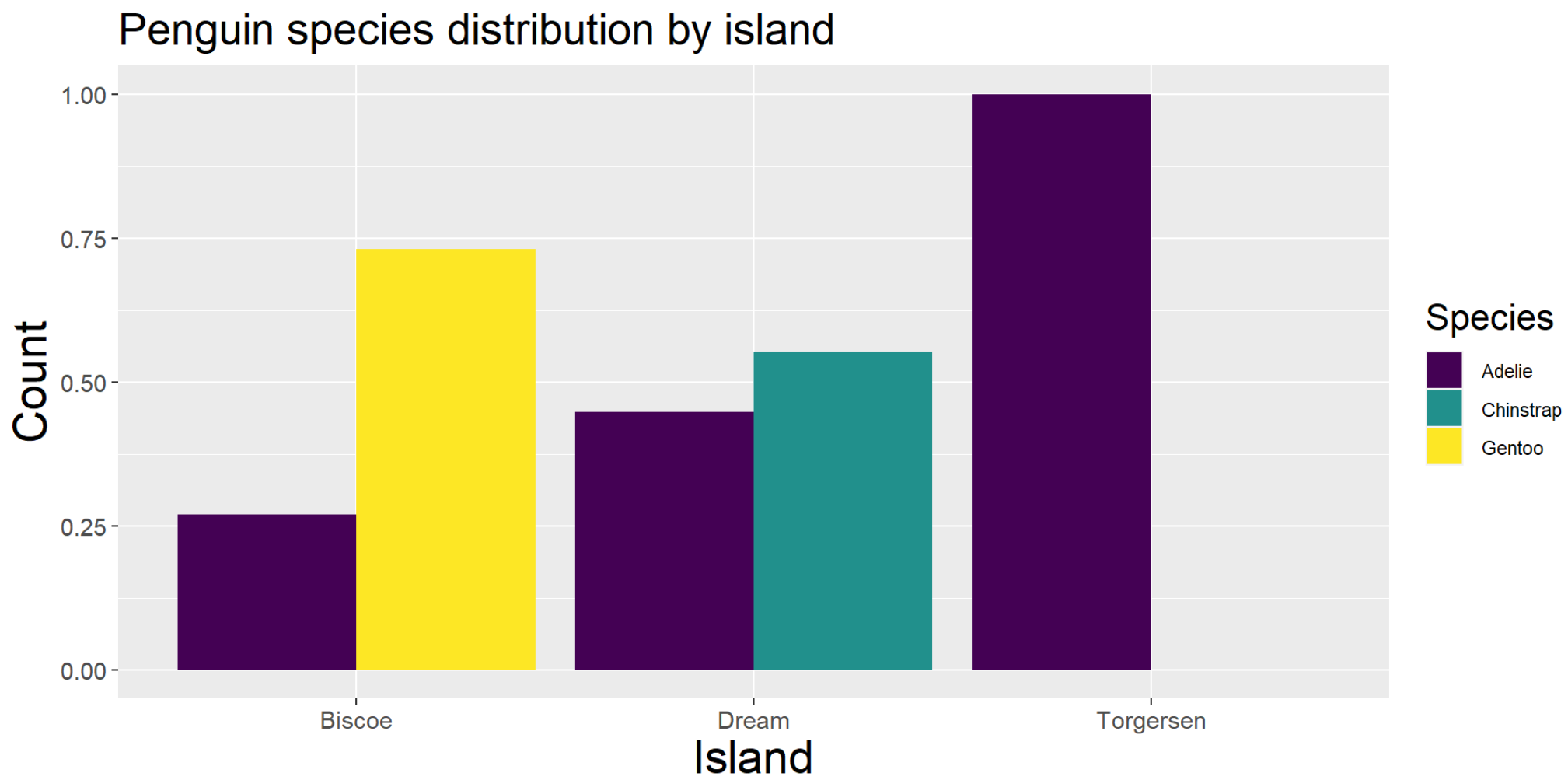
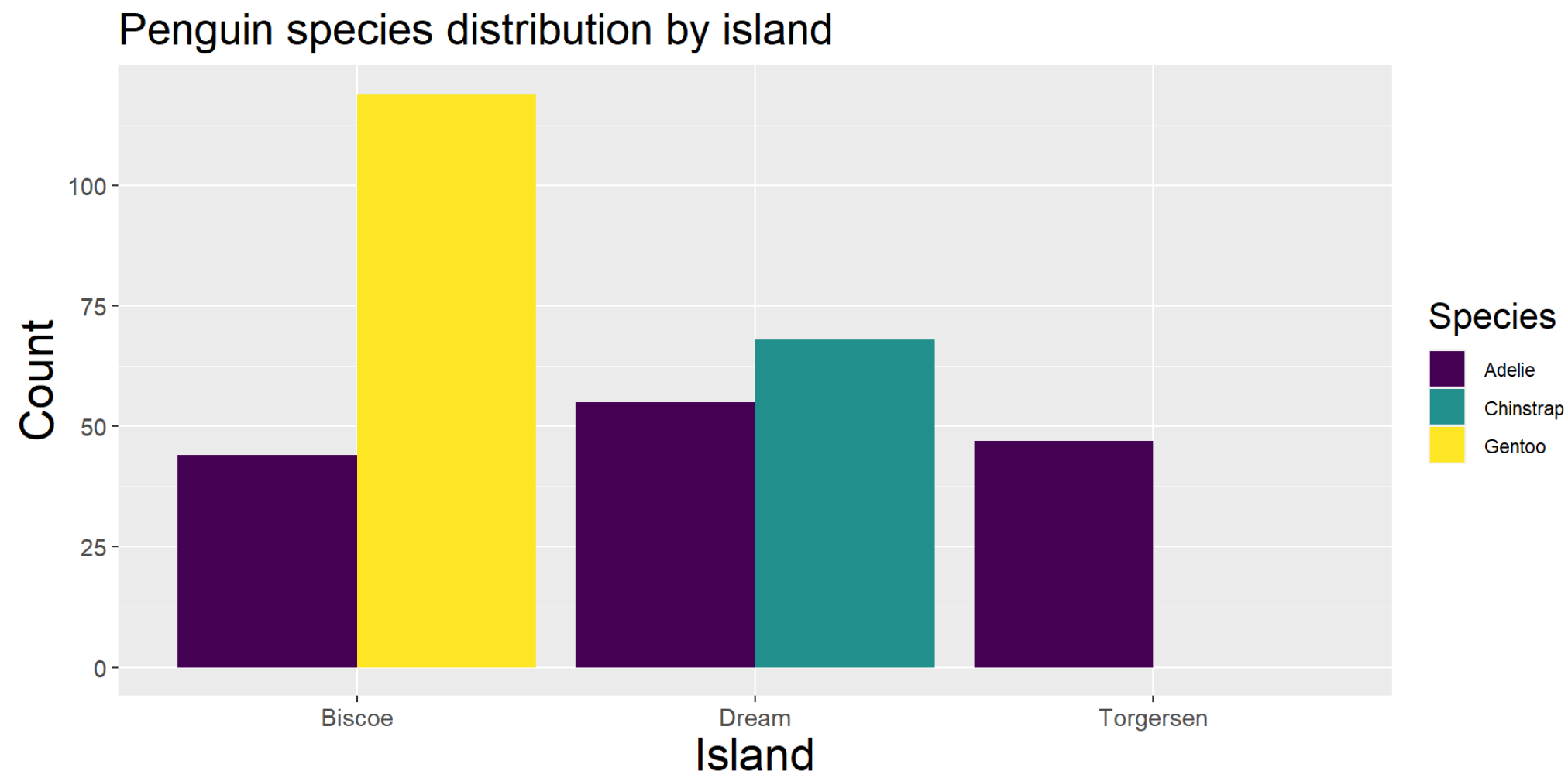
1 Numerical Variable: *Histogram*



1 Categorical Variable: *Bar Chart*

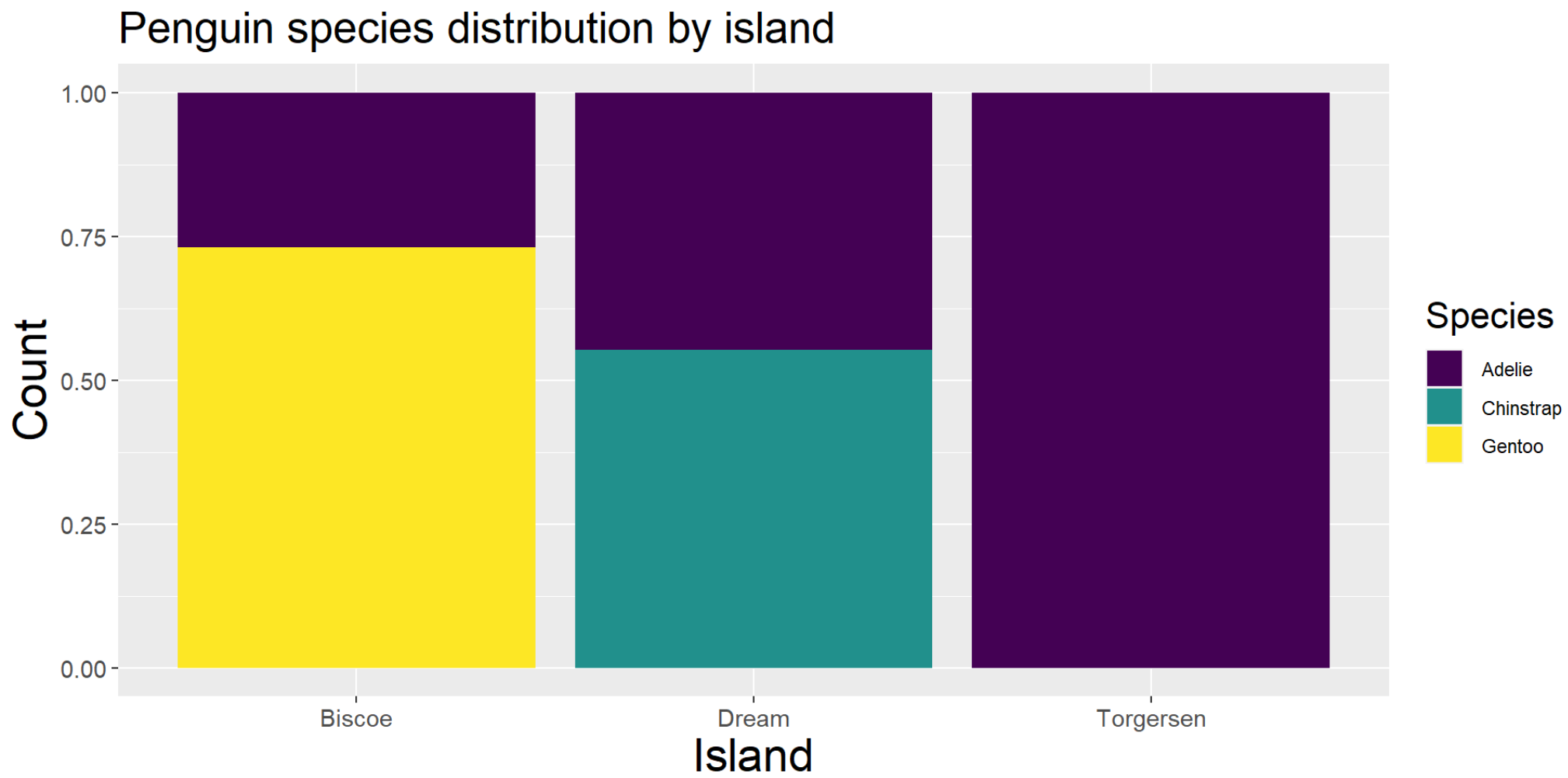
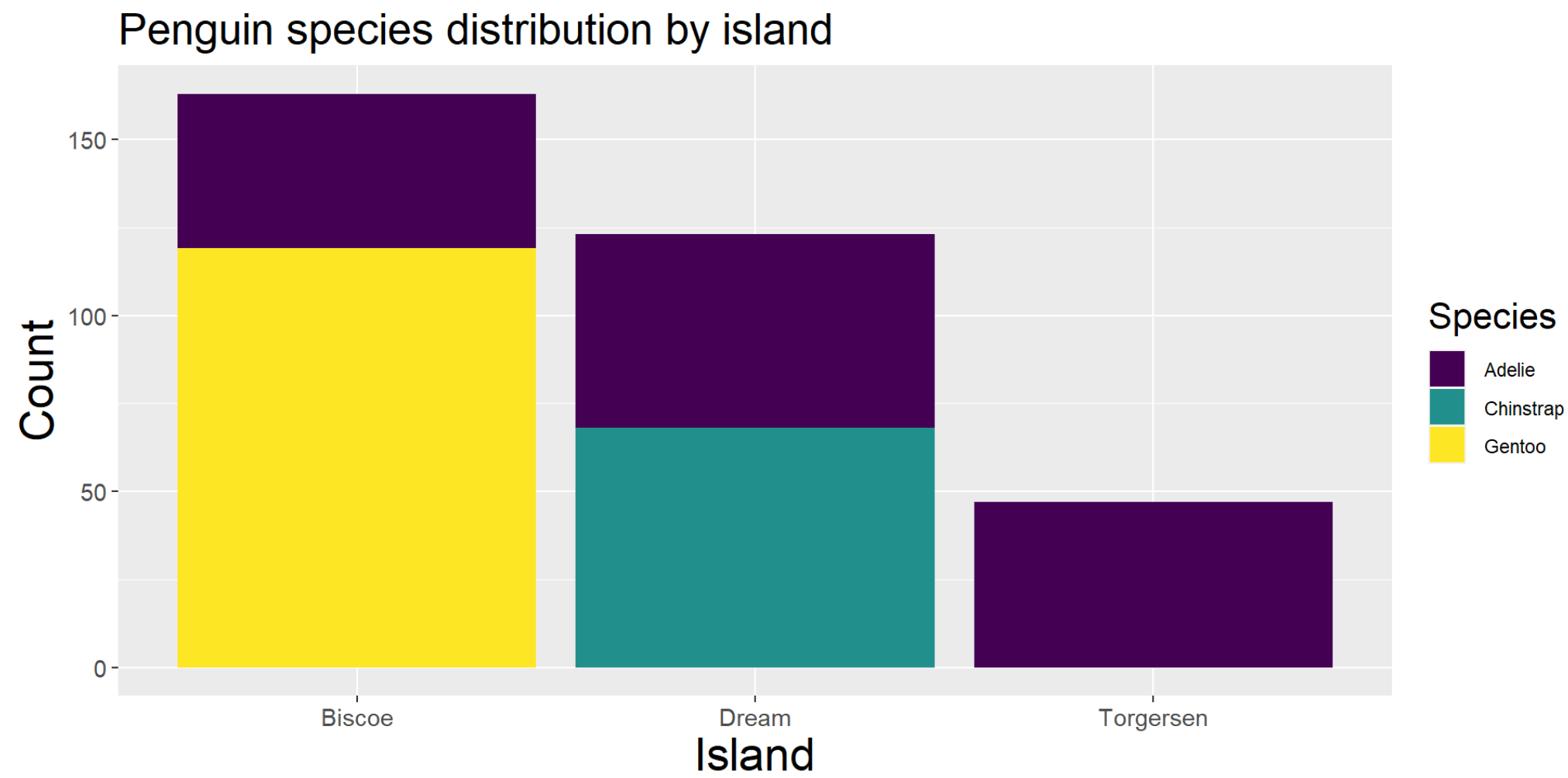
2 Categorical Variables:

Side-by-side bar chart



2 Categorical Variables:

Stacked bar chart

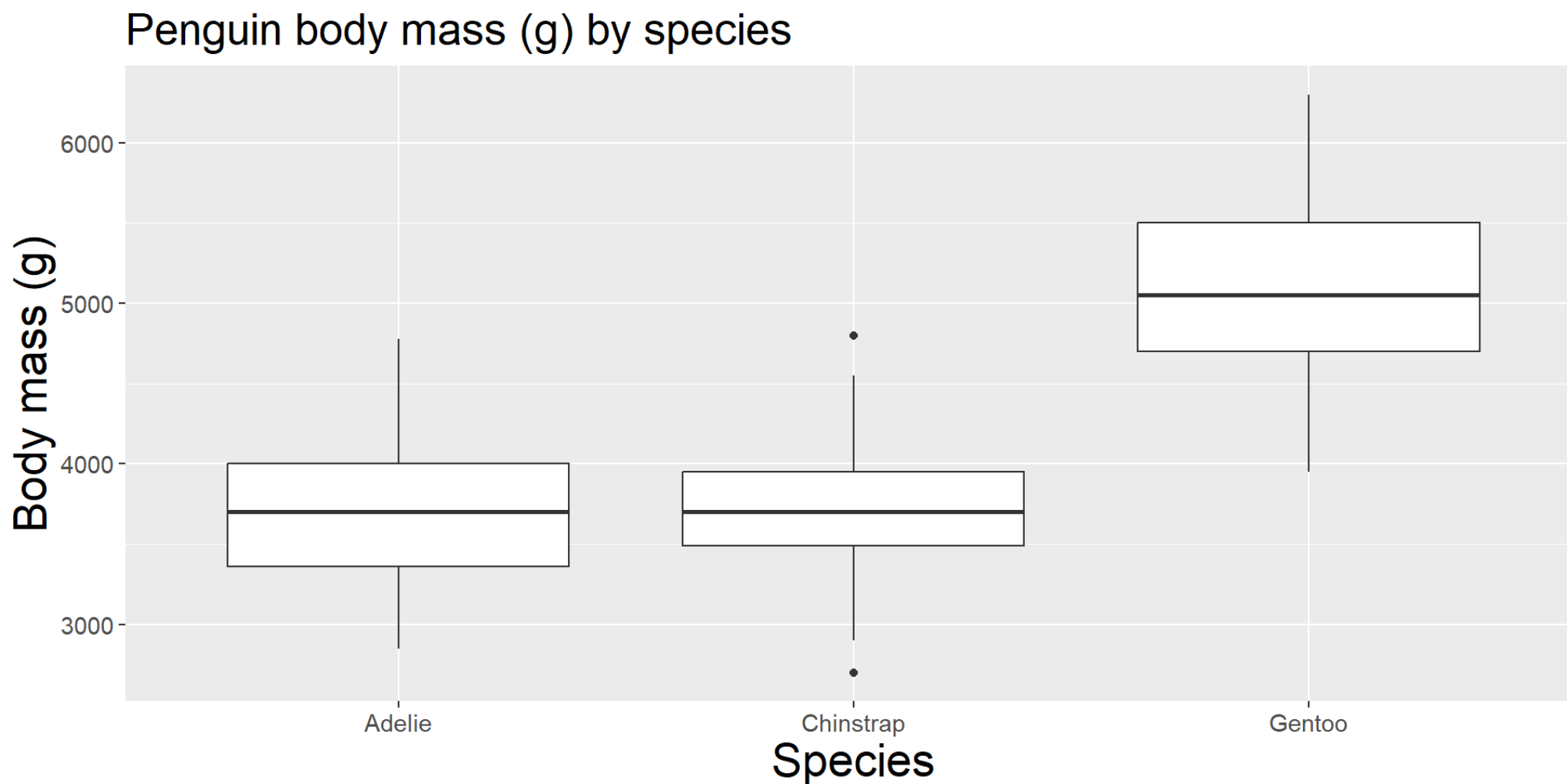
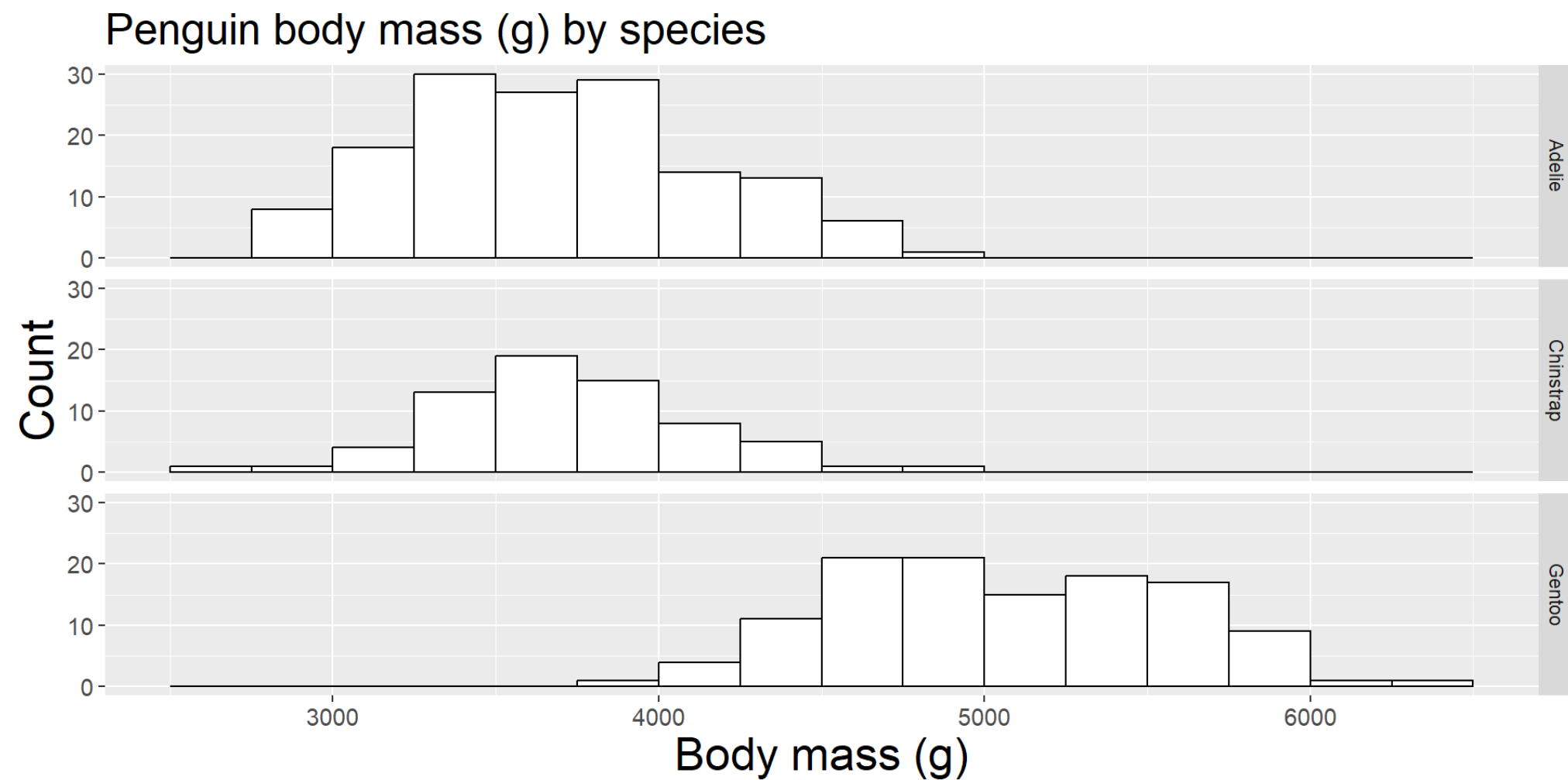


1 Categorical Variable

1 Numerical Variable

Side-by-side boxplot (bottom)

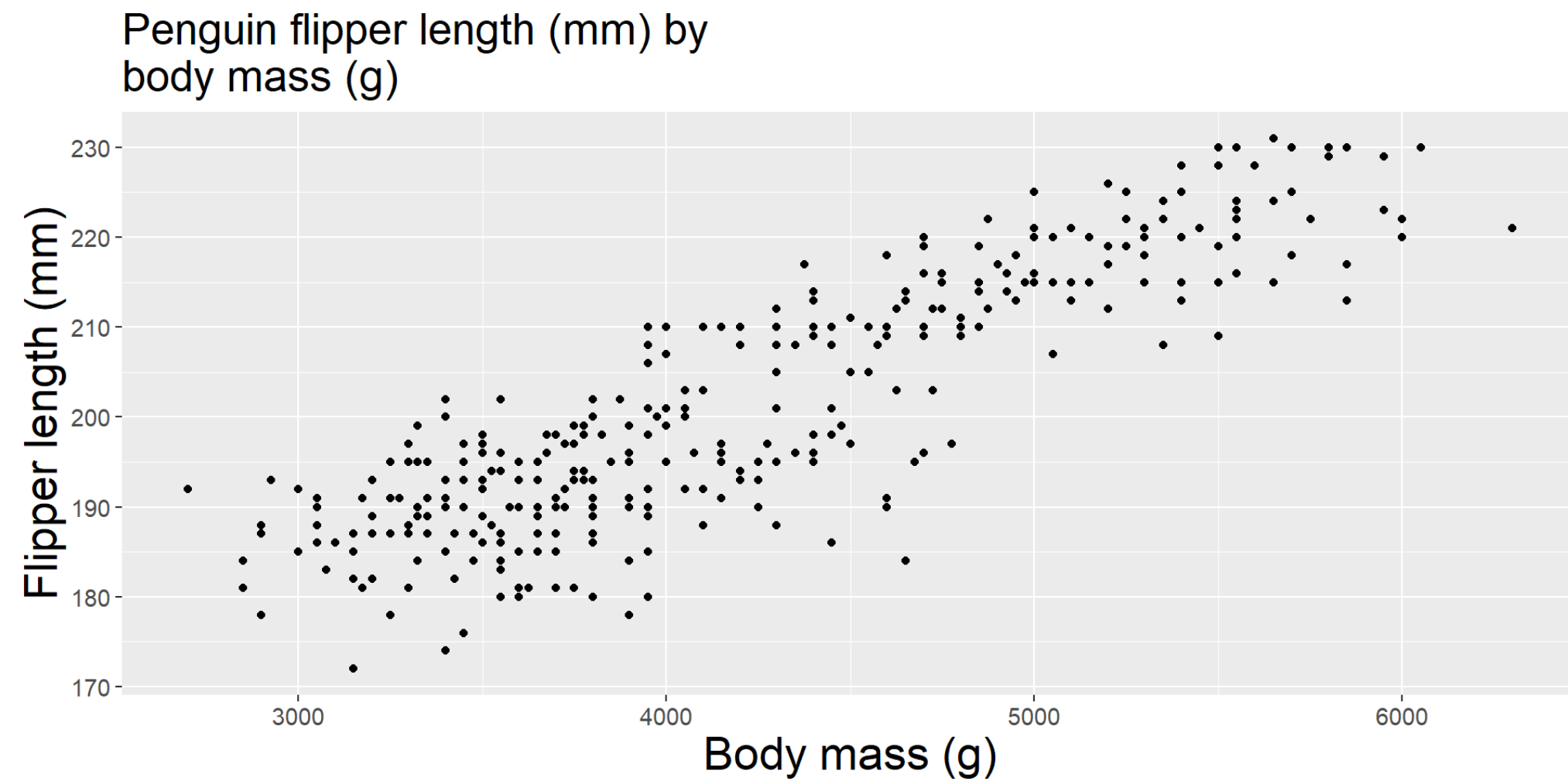
Stacked histogram (right)



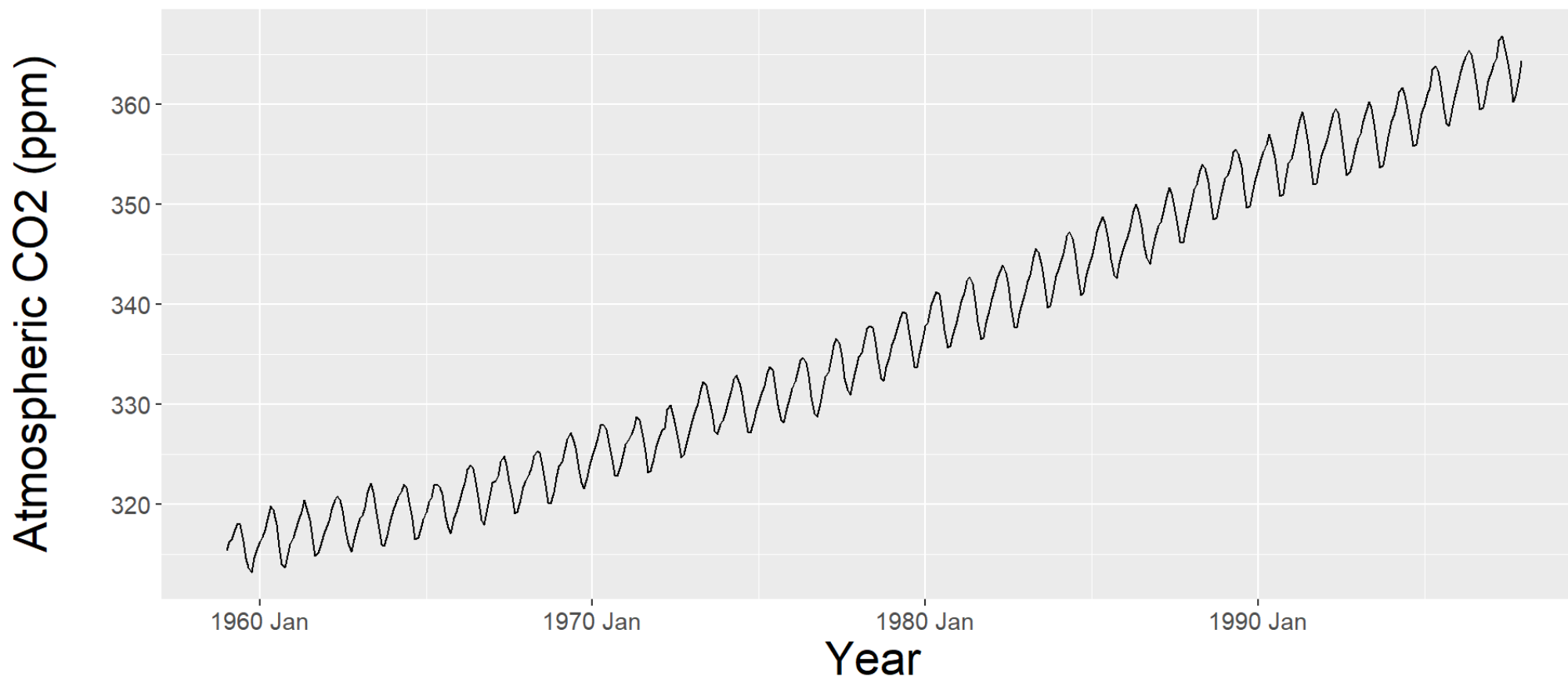
2 Numerical Variables

Scatterplot (right)

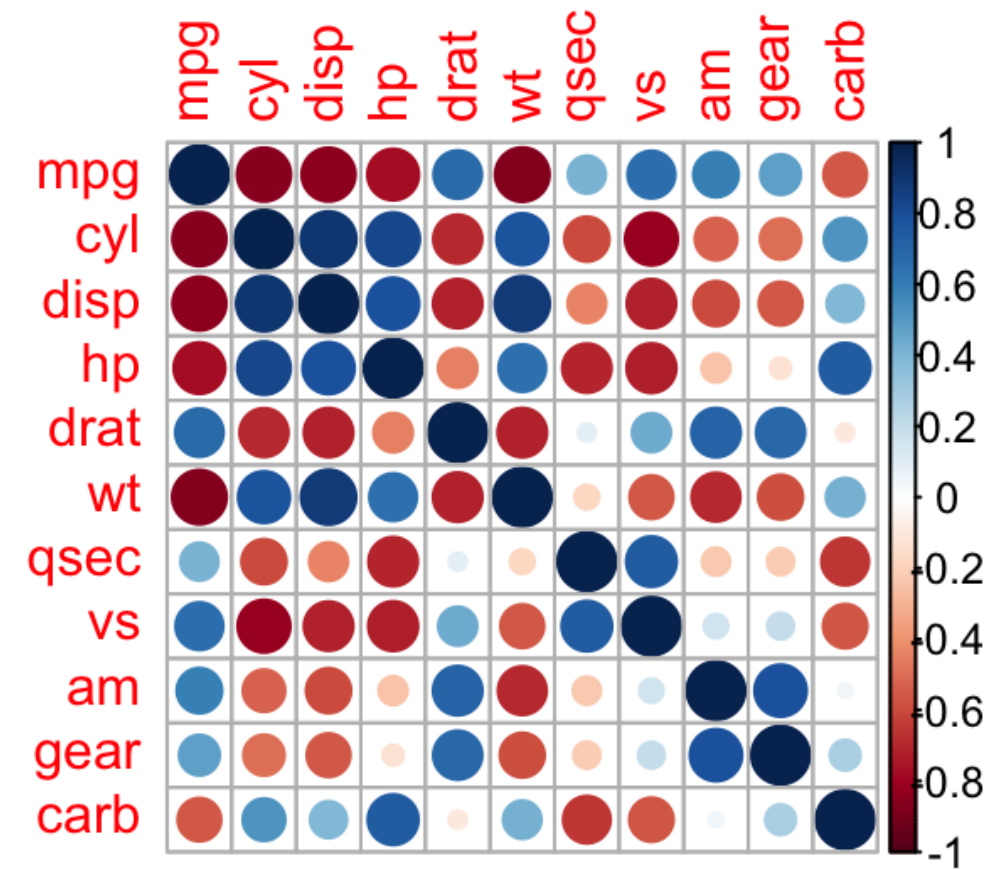
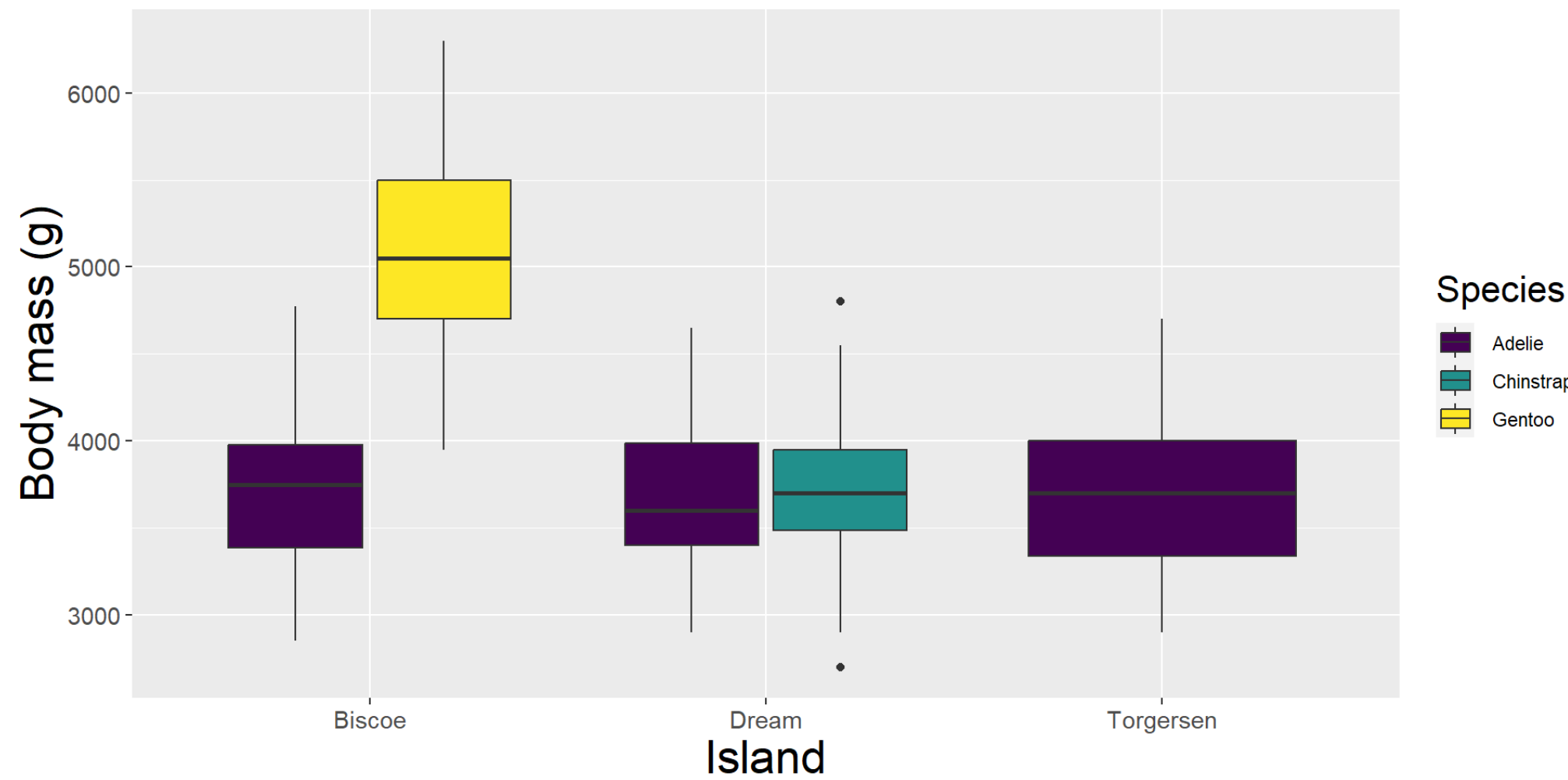
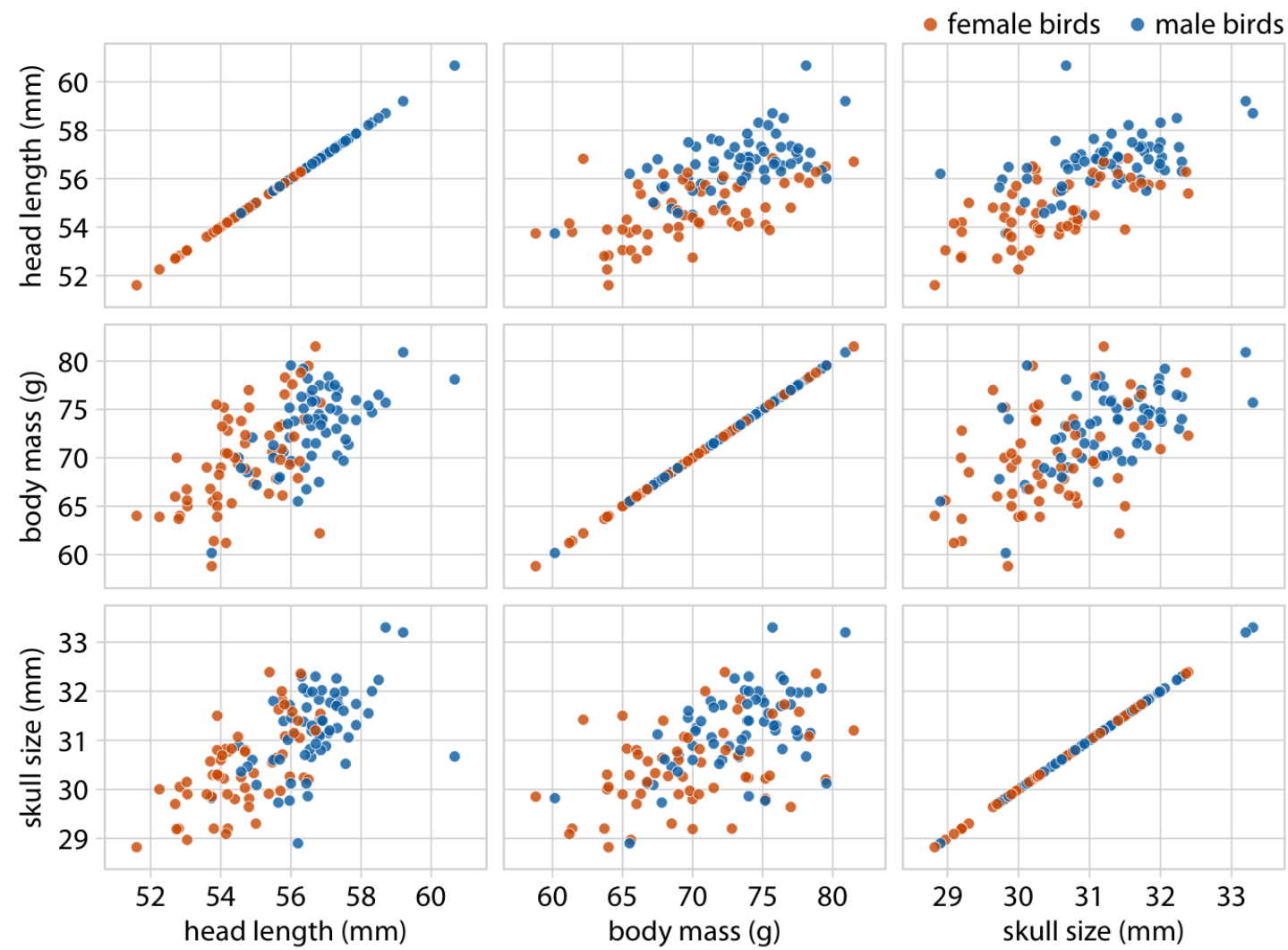
Lineplot (bottom)



Atmospheric CO2 in ppm (1959 to 1998)
Mauna Loa

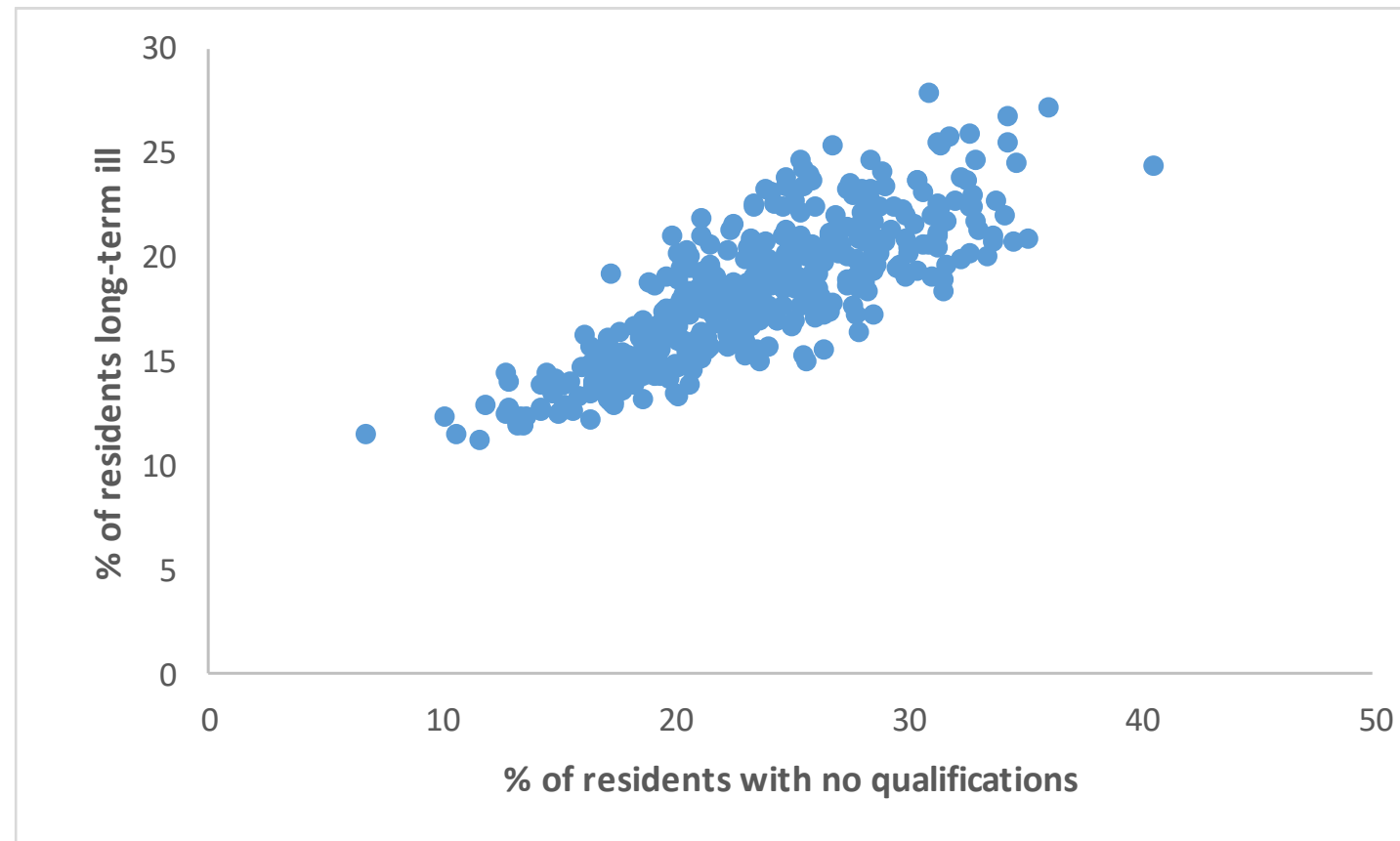


3+ Variables

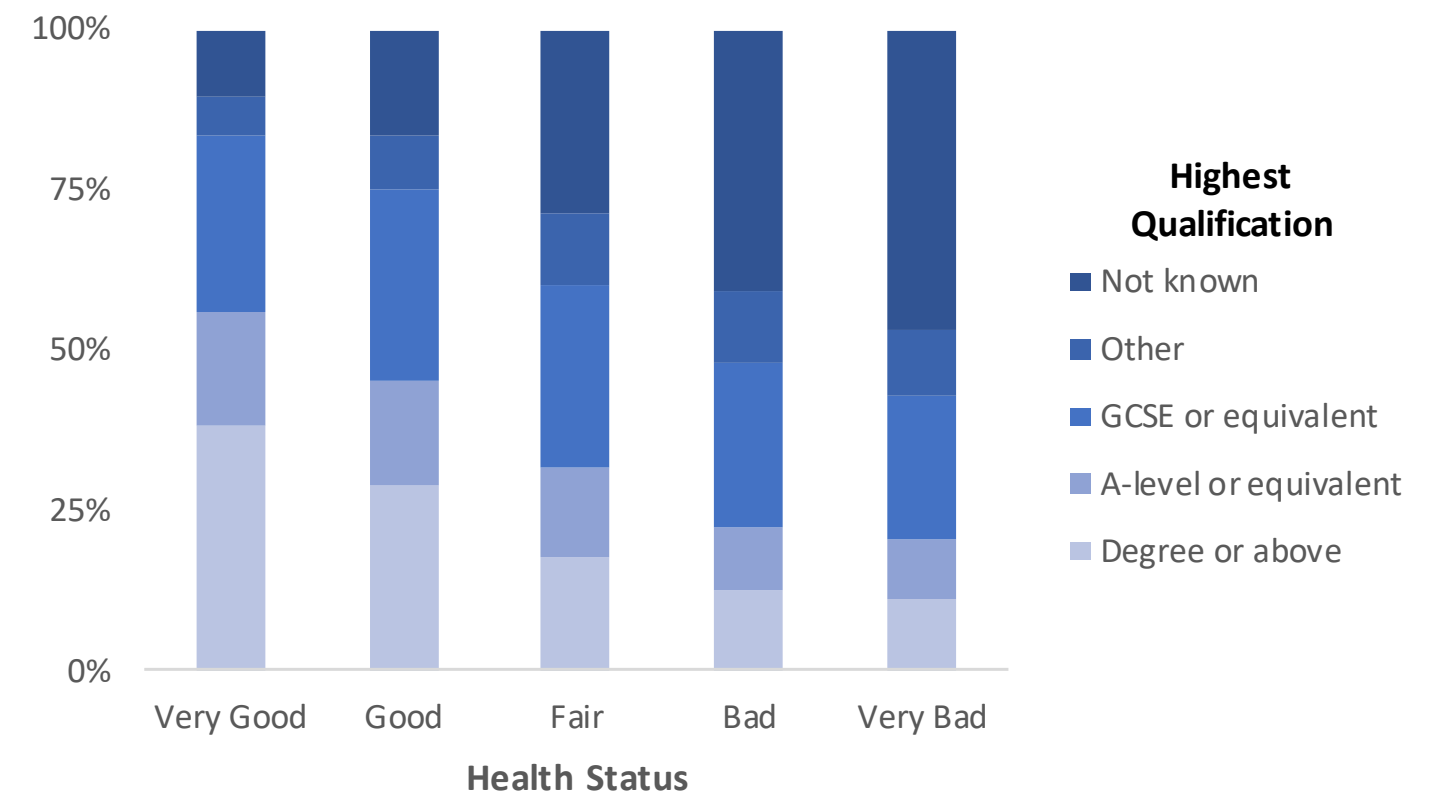


Visualising Relationships

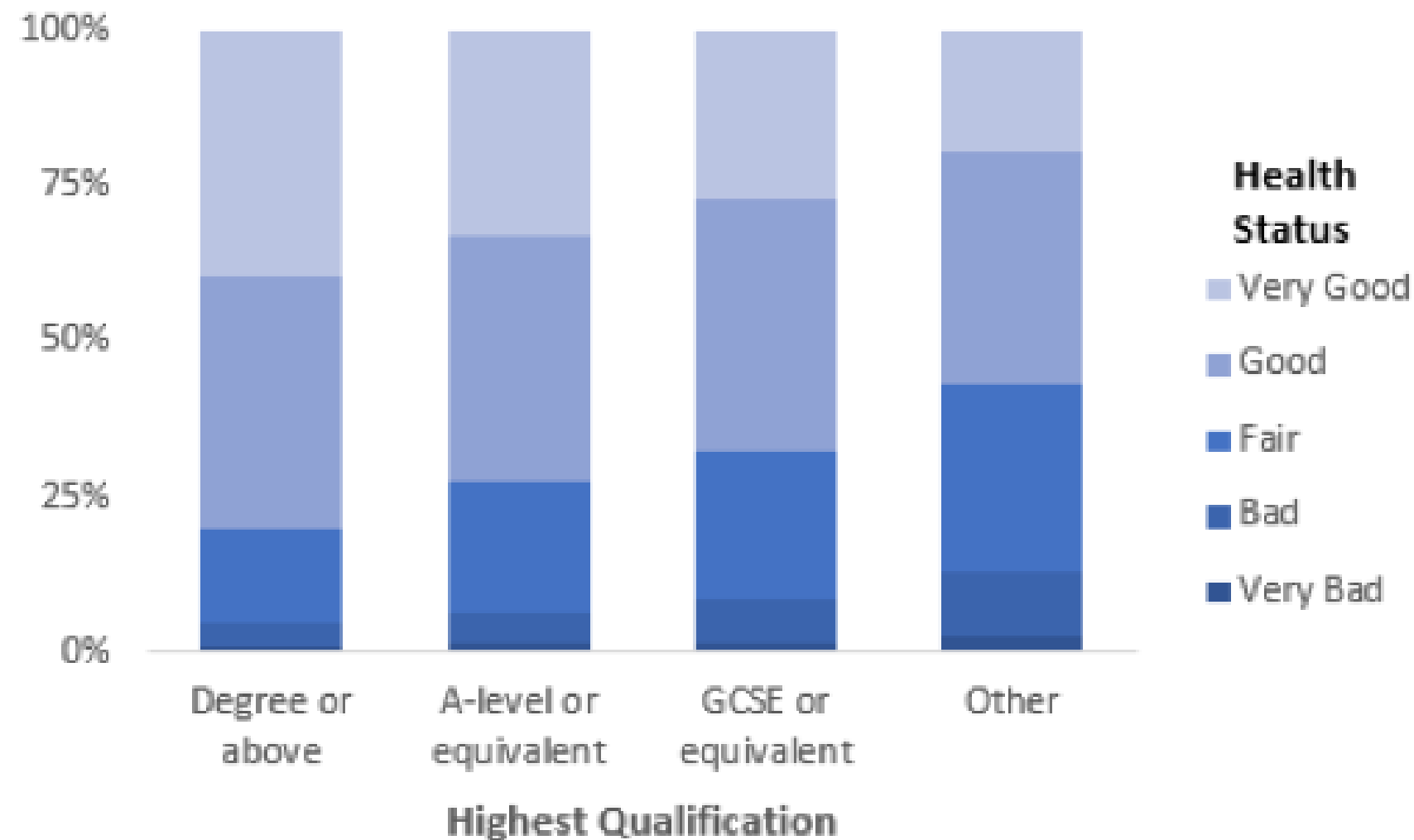
*For continuous data:
Scatter-plots*



*For categorical data:
Stacked percentage bar-charts*



Key Elements



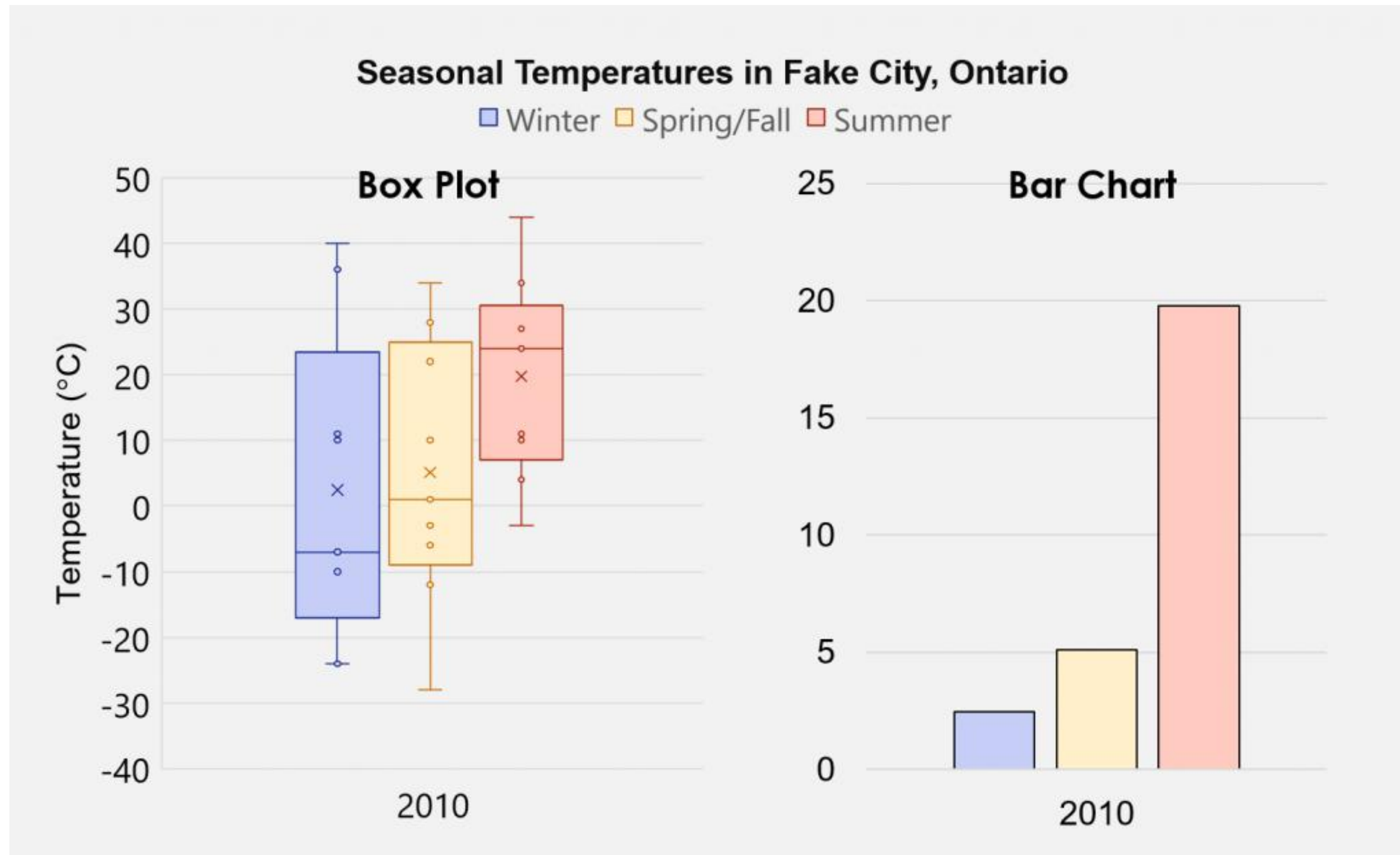
- Title (Descriptive)
- Number
- Data source (if necessary)
- Axis labels & titles
- Legend
- Clear colours

Figure 1 Health and Highest Qualification

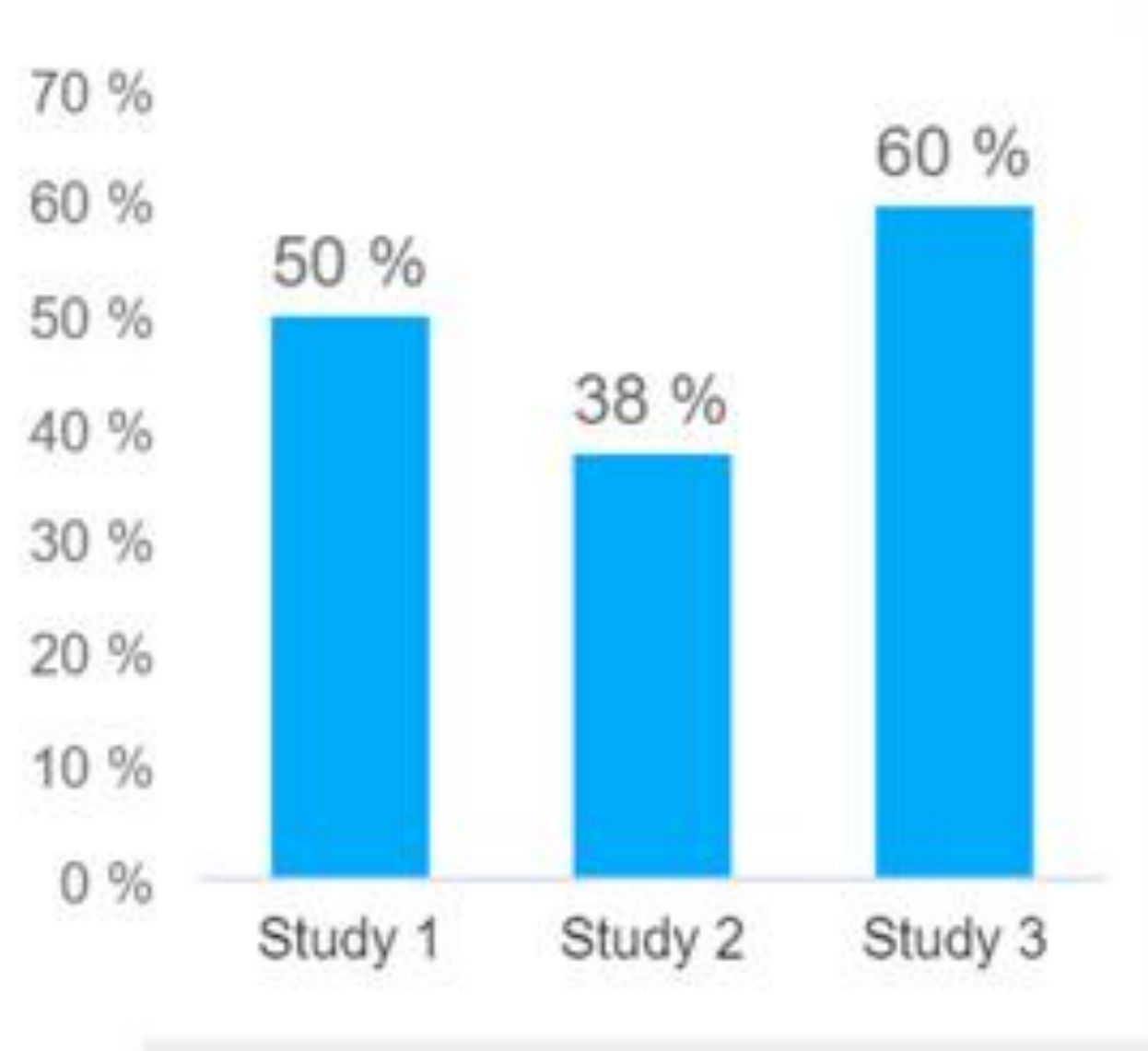
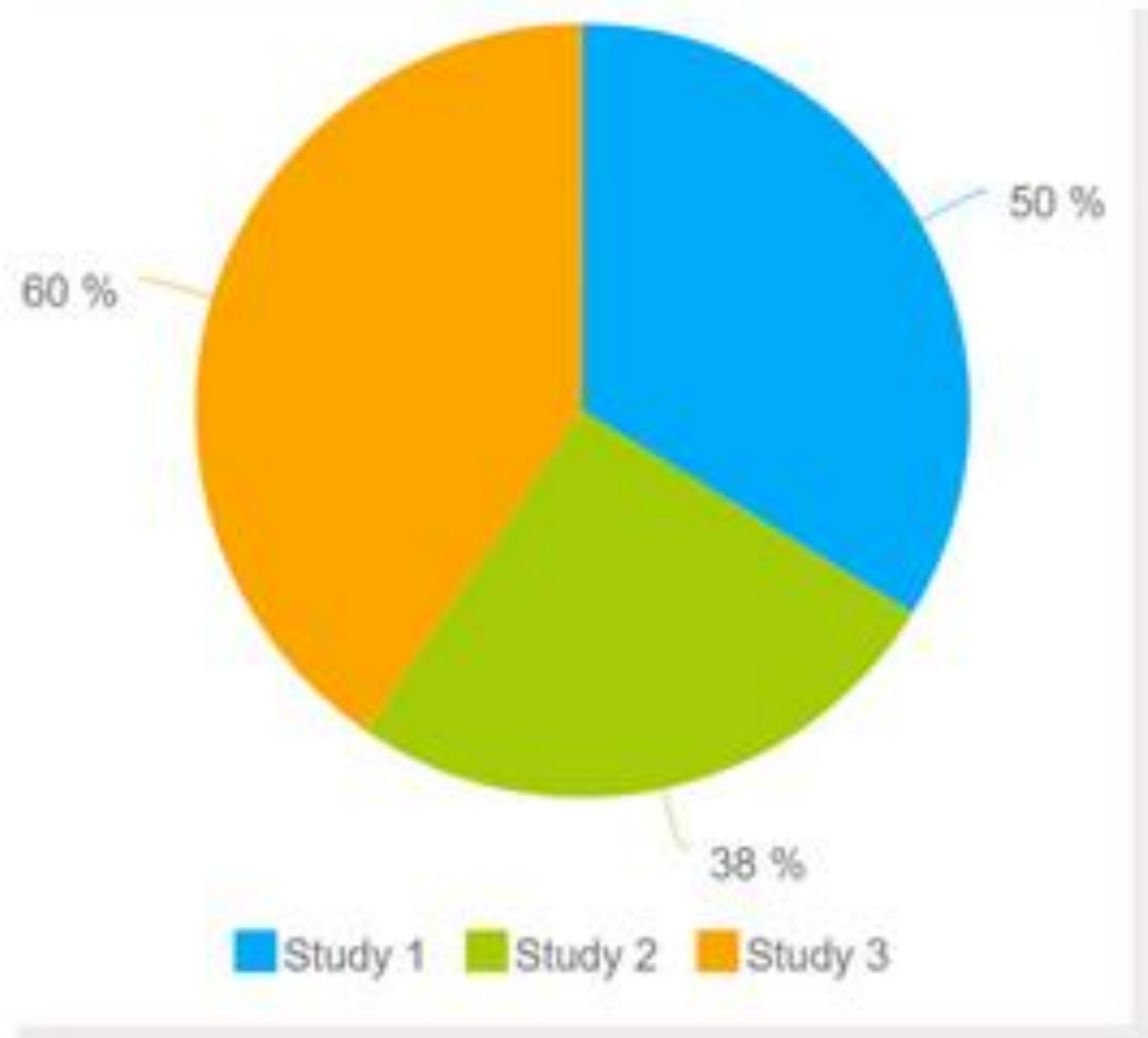
Source: Author's calculations based upon the *Family Resources Survey 2016-17* (DWP/NCSR/ONS, 2018).

Bad and Good

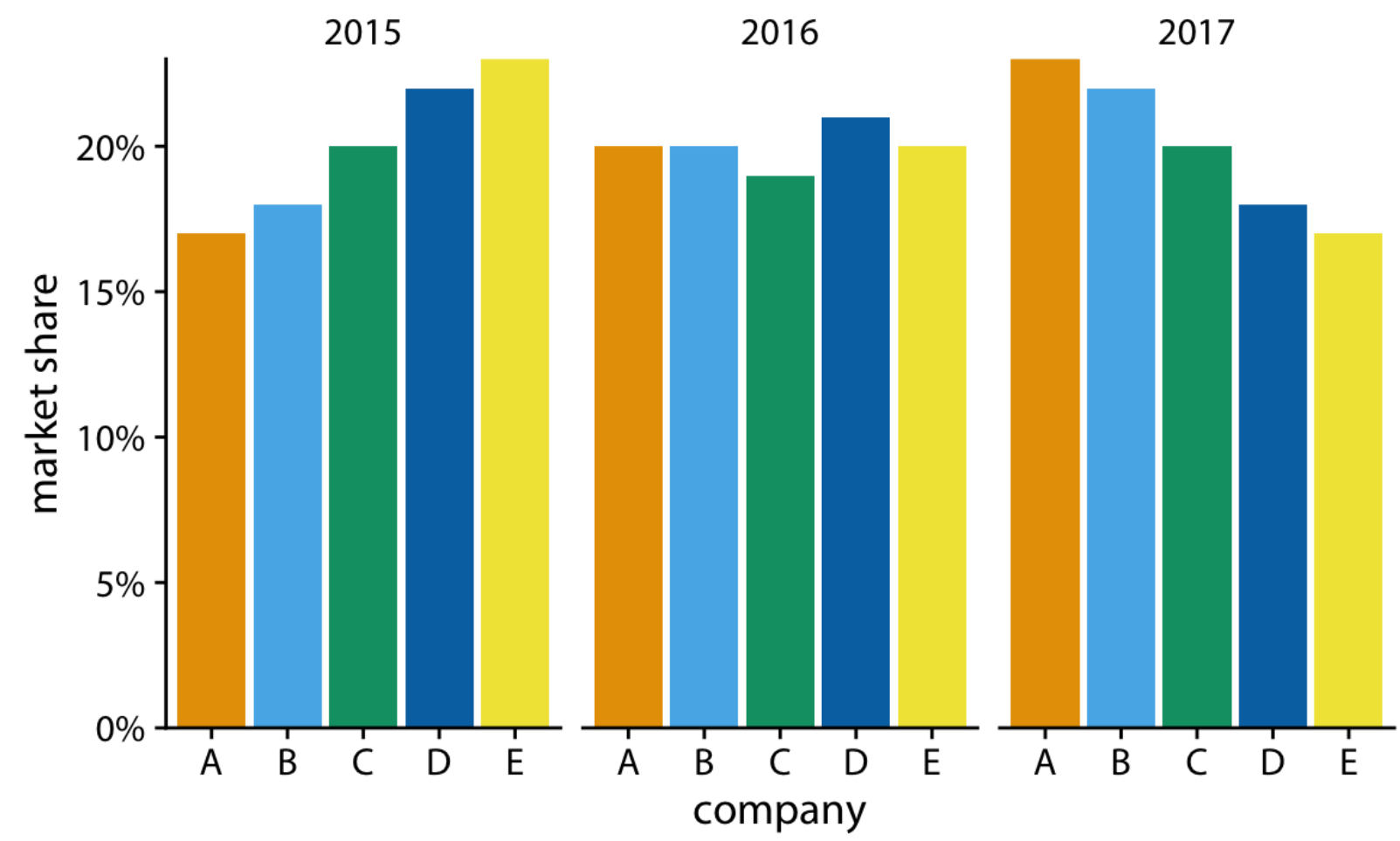
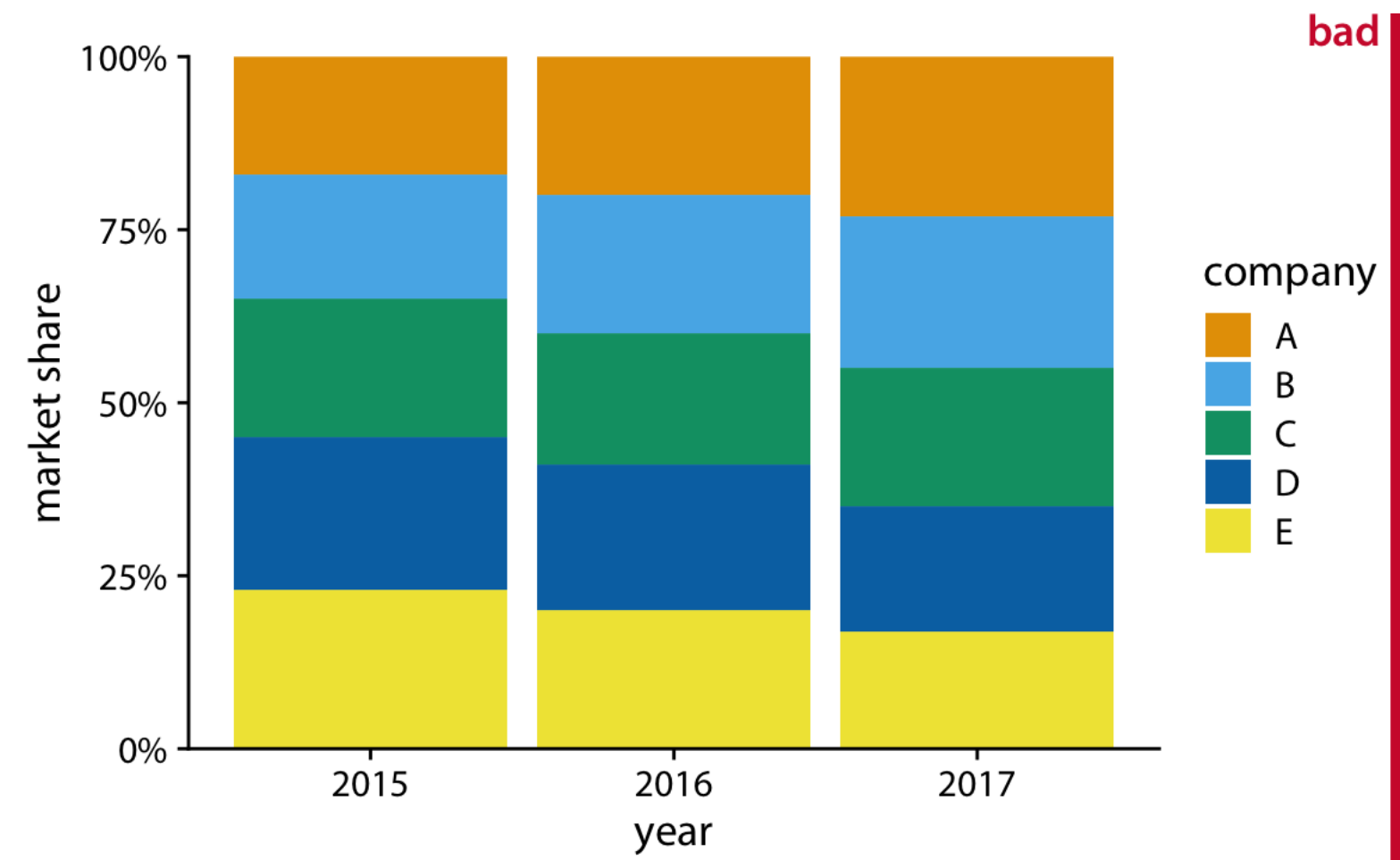
Representing Distributions



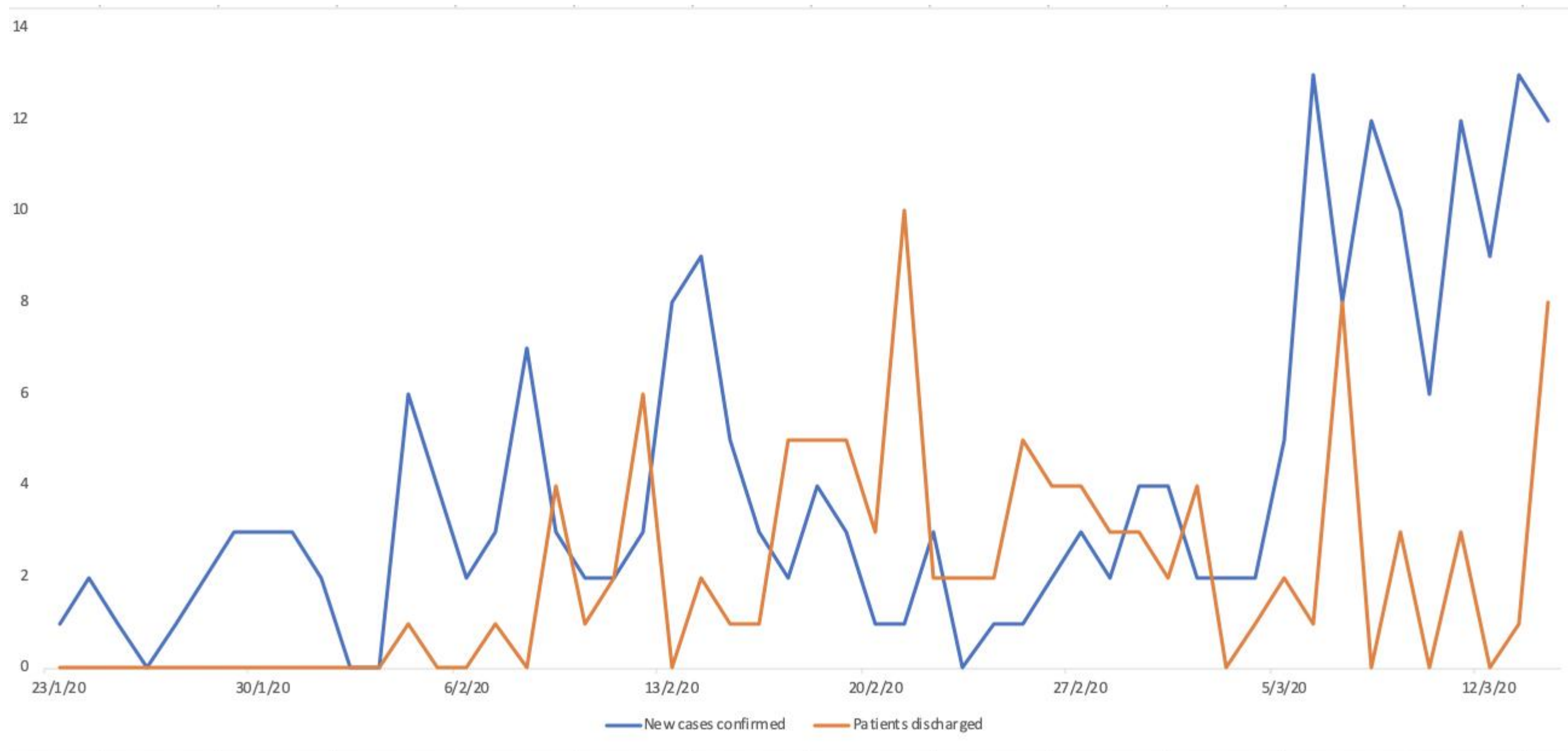
Percentages/ Proportions



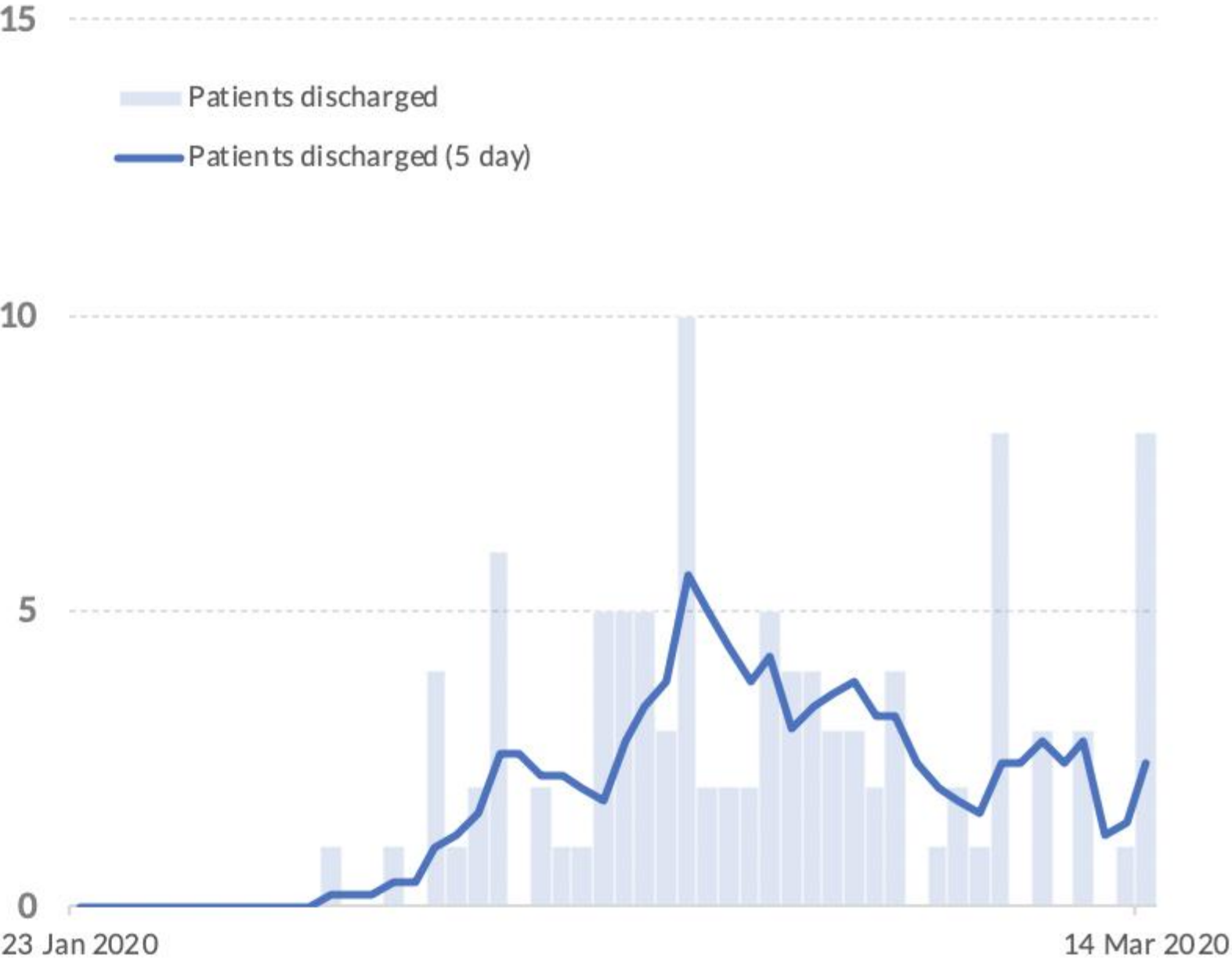
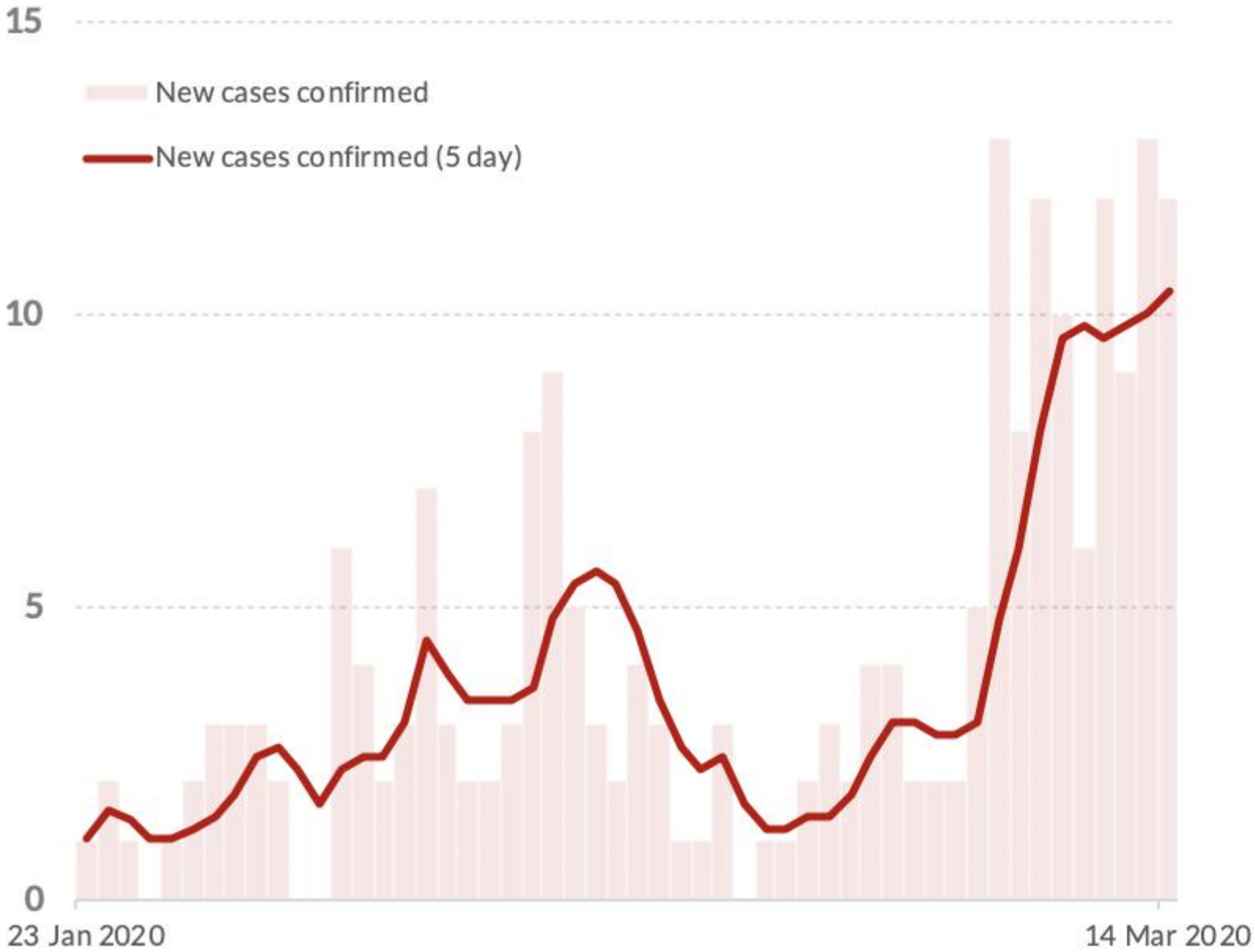
Percentages/ Proportions



Lineplots?



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



Tables

a

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

ugly

b

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

ugly



c

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

d

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