Figure prototyping with ggplot2

Lyuba V. Bozhilova

MBU PostDoc Workshop, 5 May 2022

Housekeeping notes

- You need a laptop with R and RStudio installed.
- Workshop materials: https://tinyurl.com/pfc-datavis
- Please ask questions & give feedback anytime!
- Irrelevant, but some of us run on Thursdays…

Outline

- Part I: What is the Grammar of Graphics?
- Part II: Creating complex, versatile plots
- Part III: Refining figures

There are exercises throughout, go at your own pace.

Already using ggplot2? Try #TidyTuesday.

Outline

- Part I: What is the Grammar of Graphics? ⇨ **Slides, now**
- Part II: Creating complex, versatile plots
- Part III: Refining figures

} **Code, later**

There are exercises throughout, go at your own pace.

Already using ggplot2? Try #TidyTuesday.

The Grammar of Graphics

Big Idea: Plots are made by adding separate, (semi)-independent elements together.

Some elements are essential, they make the plot work:

- data
- aesthetic mappings
- geometric objects

} **Part II**

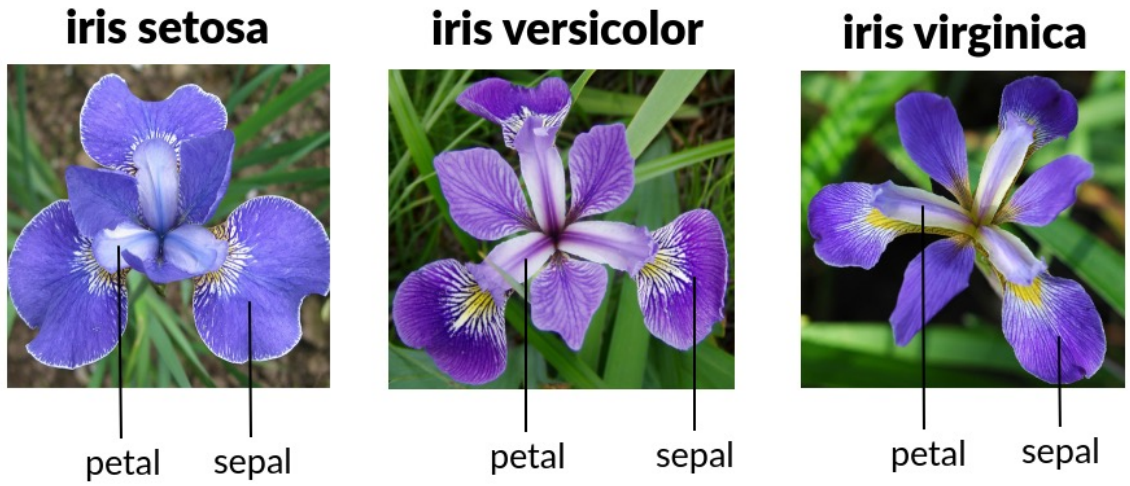
Some elements are there to make things pretty:

- scales
- annotation
- themes

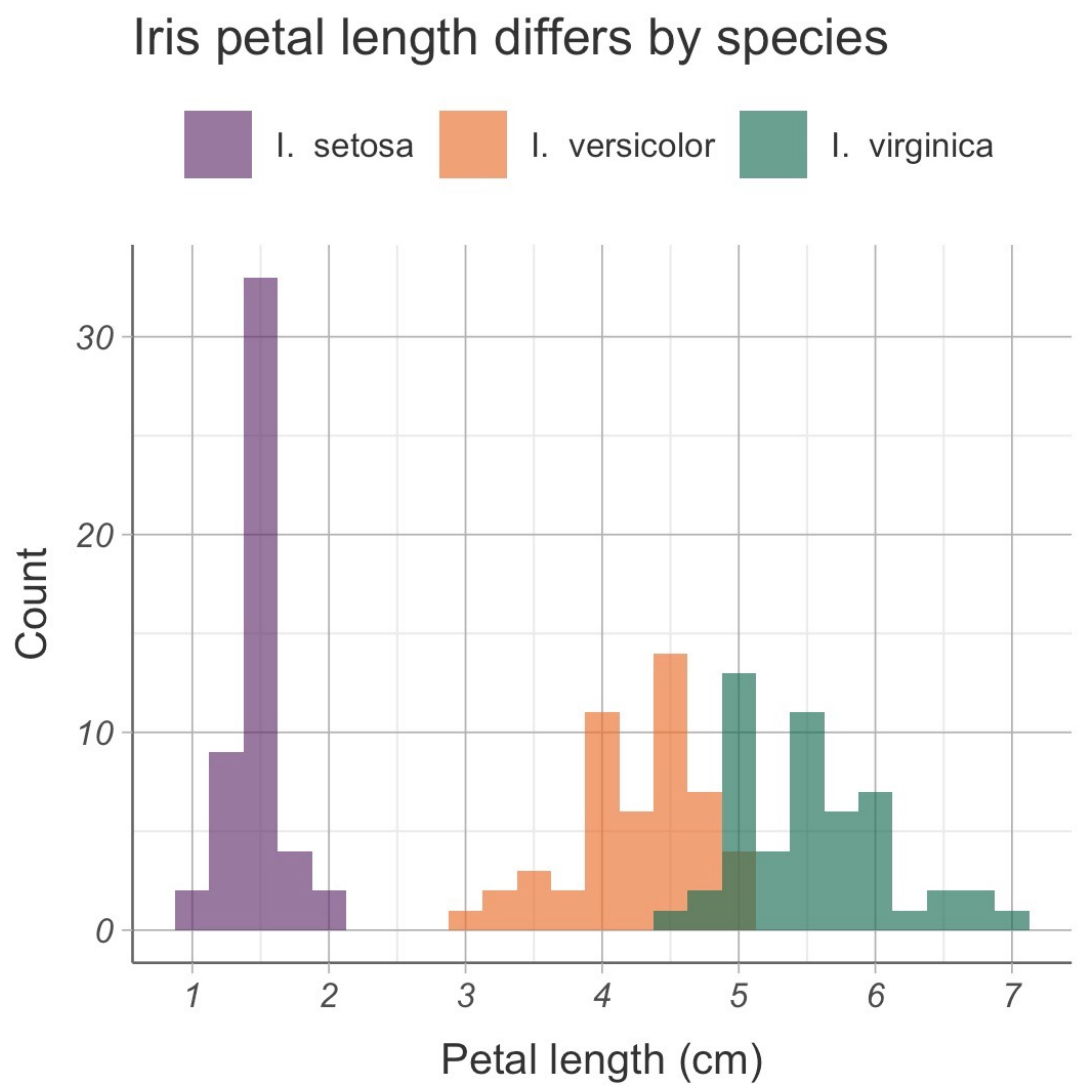
} **Part III**

Fisher's iris data

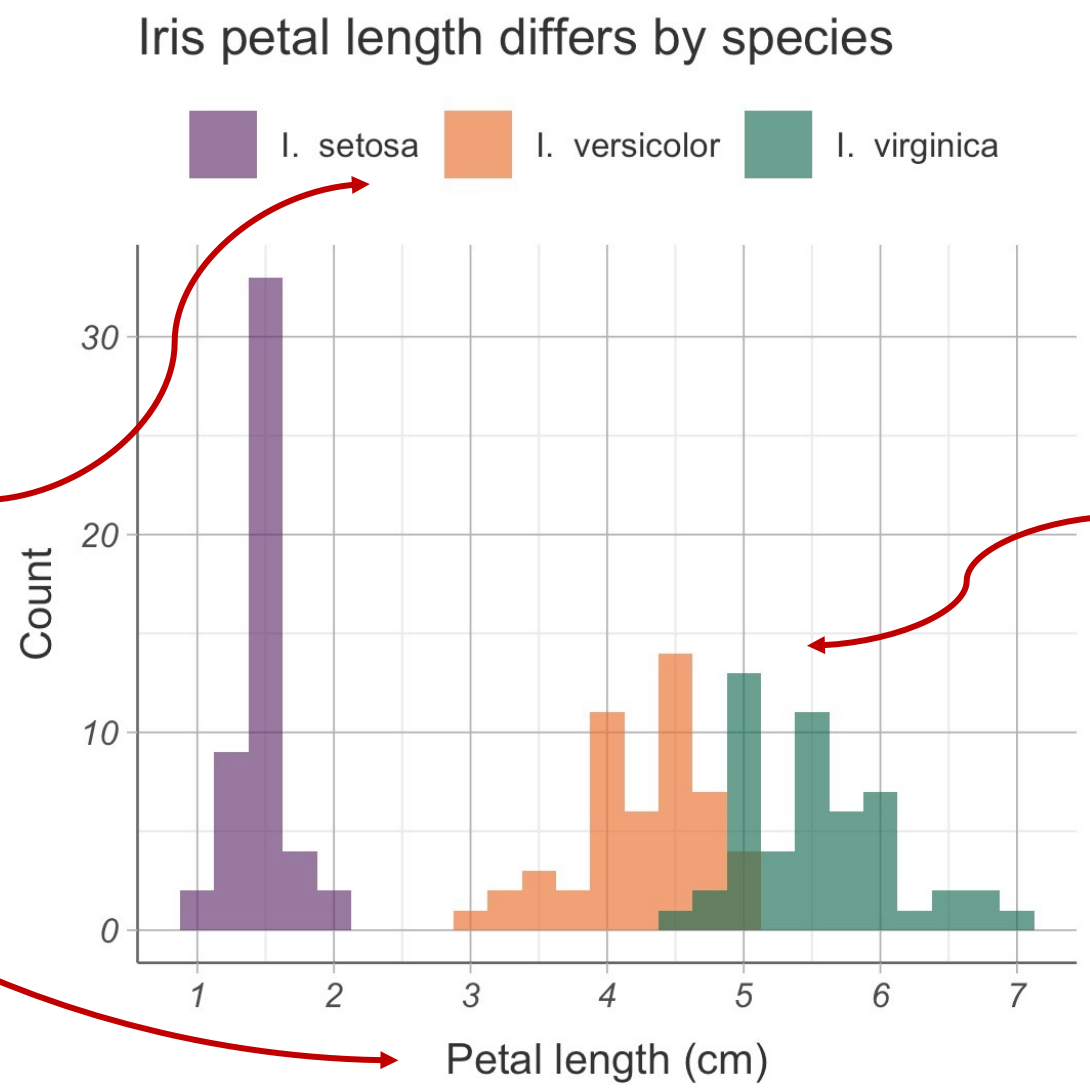
- Famous toy dataset
- Five variables, 150 observations
- Data frame format
  - columns are variables
  - rows are observations




```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
## 2          4.9         3.0          1.4         0.2  setosa
## 3          4.7         3.2          1.3         0.2  setosa
## 4          4.6         3.1          1.5         0.2  setosa
## 5          5.0         3.6          1.4         0.2  setosa
## 6          5.4         3.9          1.7         0.4  setosa
```


Iris petal length differs by species
I. setosa
I. versicolor
I. virginica
Count
30
20
10
0
1
2
3
4
5
6
7
Petal length (cm)

Data: iris data



Aesthetic mappings
colour: species
x-axis: petal length

Geometric object
histogram

Core ggplot2 elements

Data

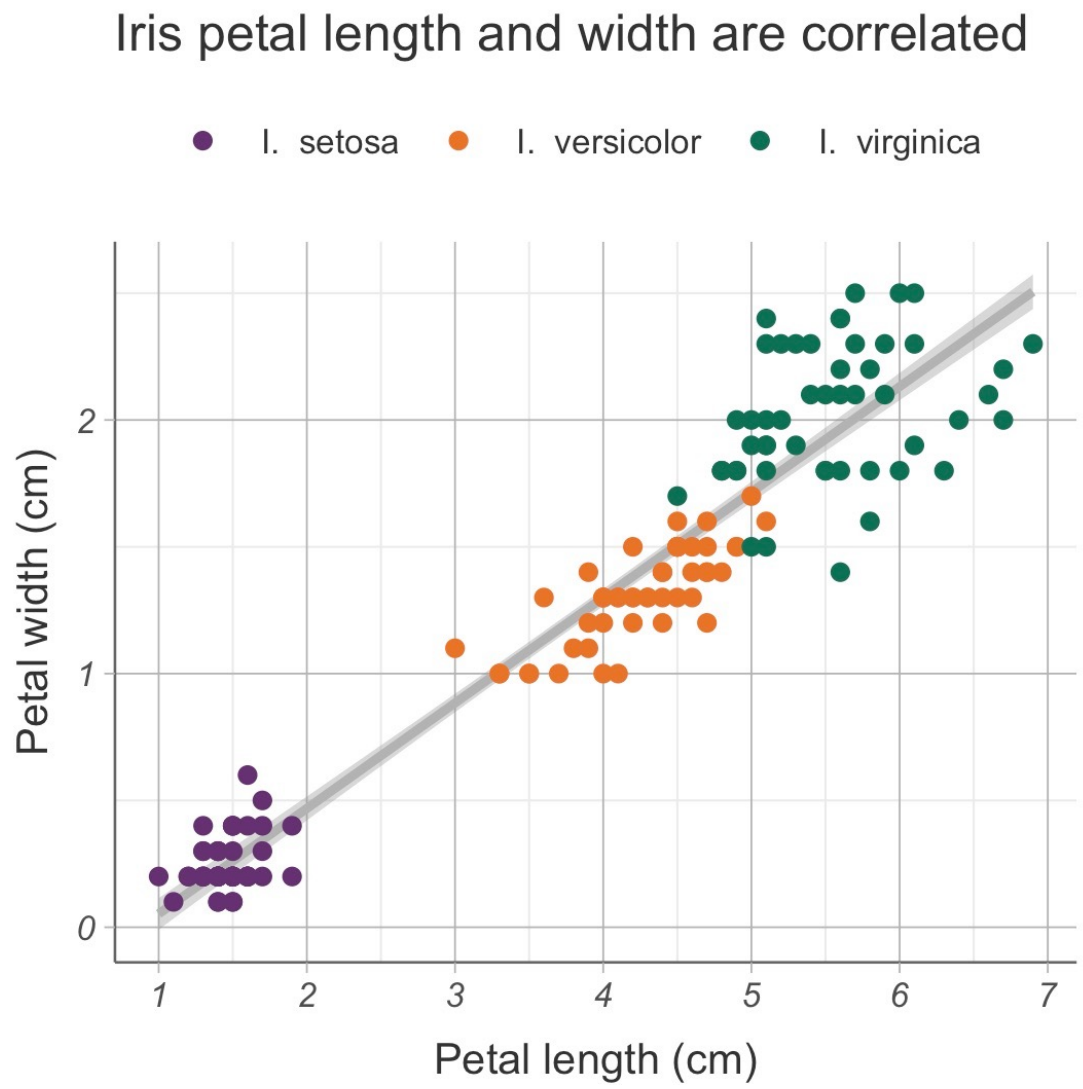
- Multiple observations of one or more variables
- Variables can be of many types: numeric, categorical, time, etc.

Aesthetic mappings

- How do the variables map onto plot elements?
- Think axes, colours, shapes, line type…

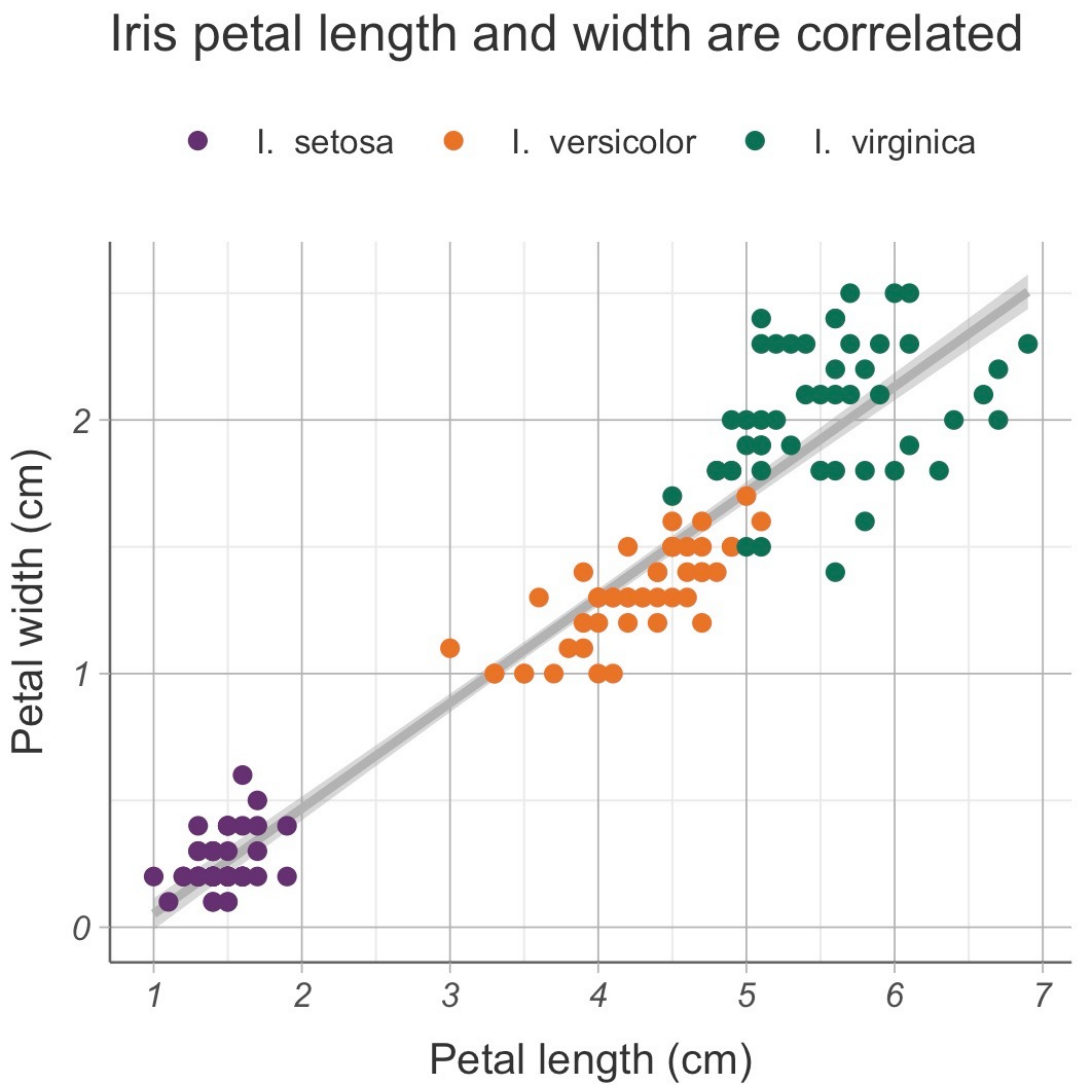
Geometric objects

- What kind of statistical plot do we see?
- E.g. histogram, scatter plot, box plot


Iris petal length and width are correlated
I. setosa
I. versicolor
I. virginica
Petal width (cm)
0
1
2
Petal length (cm)
1
2
3
4
5
6
7

Data: iris data

Aesthetic mappings
colour: species
x-axis: petal length
y-axis: petal width



Geometric objects
scatter plot
regression line

The rest of the workshop

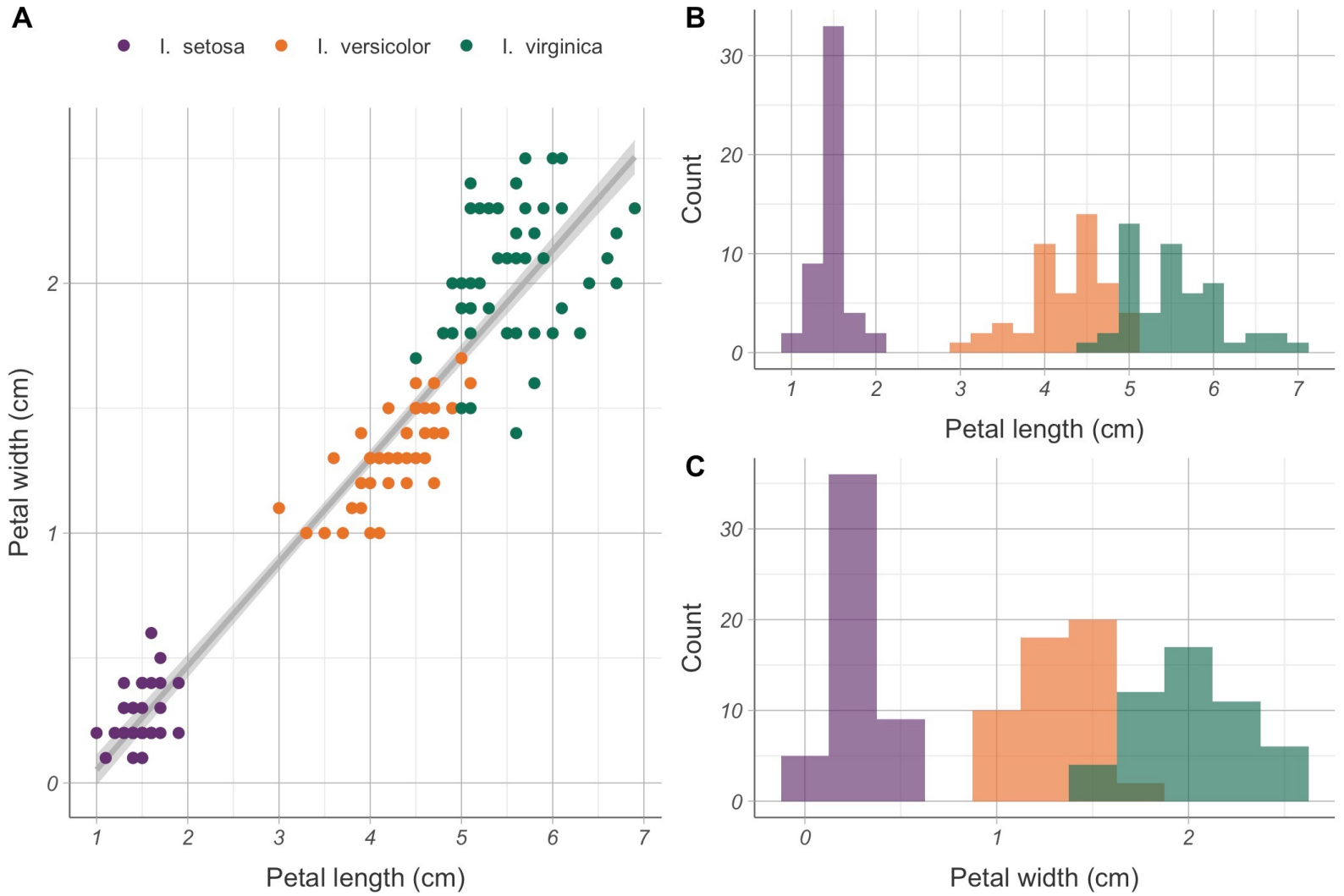
Part II: functional plots

- How ggplot2 code works
- More aesthetic mappings
- Geometric objects: how to add them together effectively, and how to use stats and positioning
- Facets

Part III: pretty figures

- Scales to make aesthetics pretty
- Themes to make everything else pretty
- Annotations can be helpful
- Make subfigures
- Save figures with good resolution and font size

End goal


A
I. setosa
I. versicolor
I. virginica
Petal width (cm)
Petal length (cm)
B
Count
Petal length (cm)
C
Count
Petal width (cm)

#TidyTuesday

Tidy Tuesday is an online community project, in which practice their tidyverse / ggplot2 skills.

- GitHub repo
- New dataset every week
  - this week: solar and wind power plants in the US
  - two weeks ago: UK cryptic crosswords
- People share their code and figures, so you can learn a lot!