

Week 6: Visualisations

POP88162 Introduction to Quantitative Research Methods

Tom Paskhalis

Department of Political Science, Trinity College Dublin

So Far

- *Vector* is the core *data structure* in R.
- Vectors can be one of the main data types (*character, numeric, logical*).
- As opposed to homogeneous vectors, *lists* allow to combine data of different types.
- *Matrices* are vectors with an added class and dimensionality attribute.
- *Data frames* are lists of equal-sized vectors.
- When subsetting data frames the techniques of subsetting matrices and lists are combined.
- *Factor variables* can be used to represent categorical data in R.

Topics for Today

- `plot()` function
- Box plot
- Good plotting practices
- `ggplot2` package

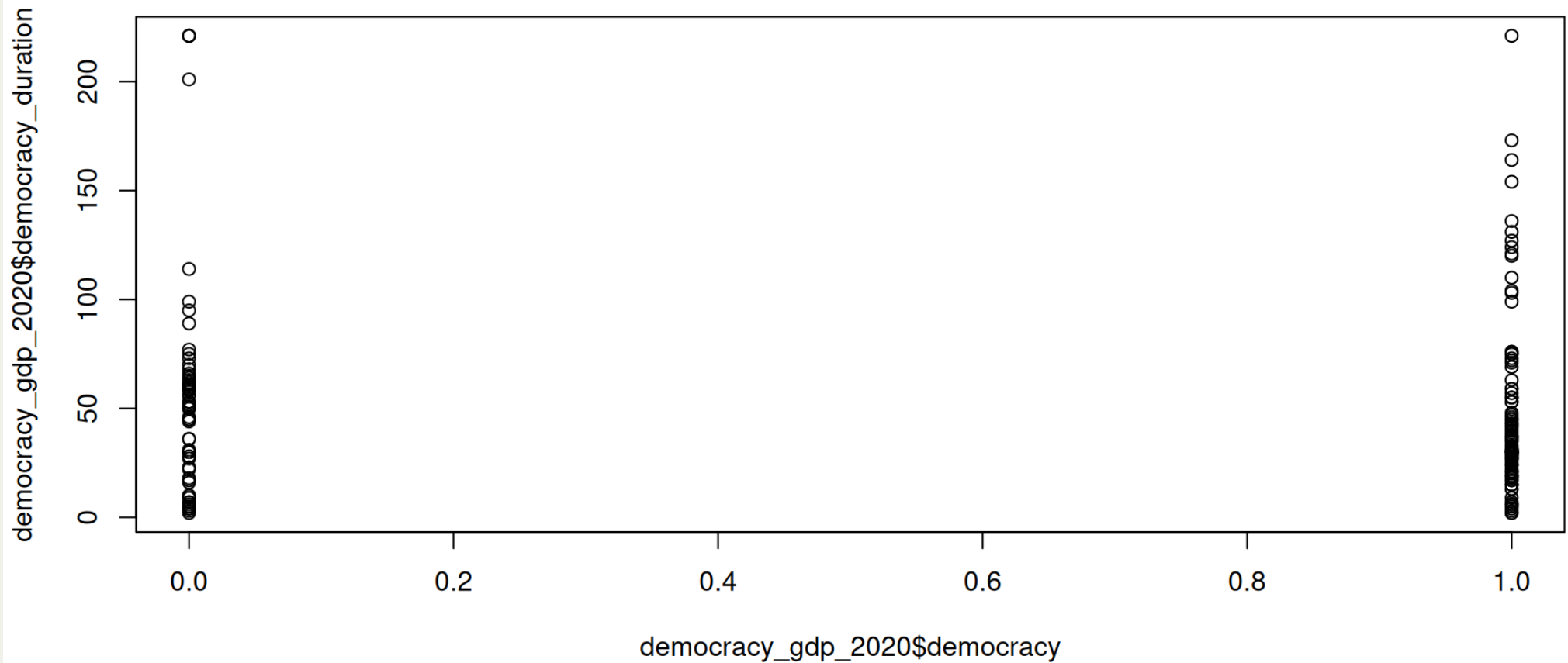
Plotting in R

- The main function for plotting in base R is `plot()`.
- It is a flexible function that relies on the data type of supplied objects.

```
1 plot(  
2   x,  
3   y,  
4   type,  
5   main,  
6   xlab,  
7   ylab  
8 )
```


Scatterplot

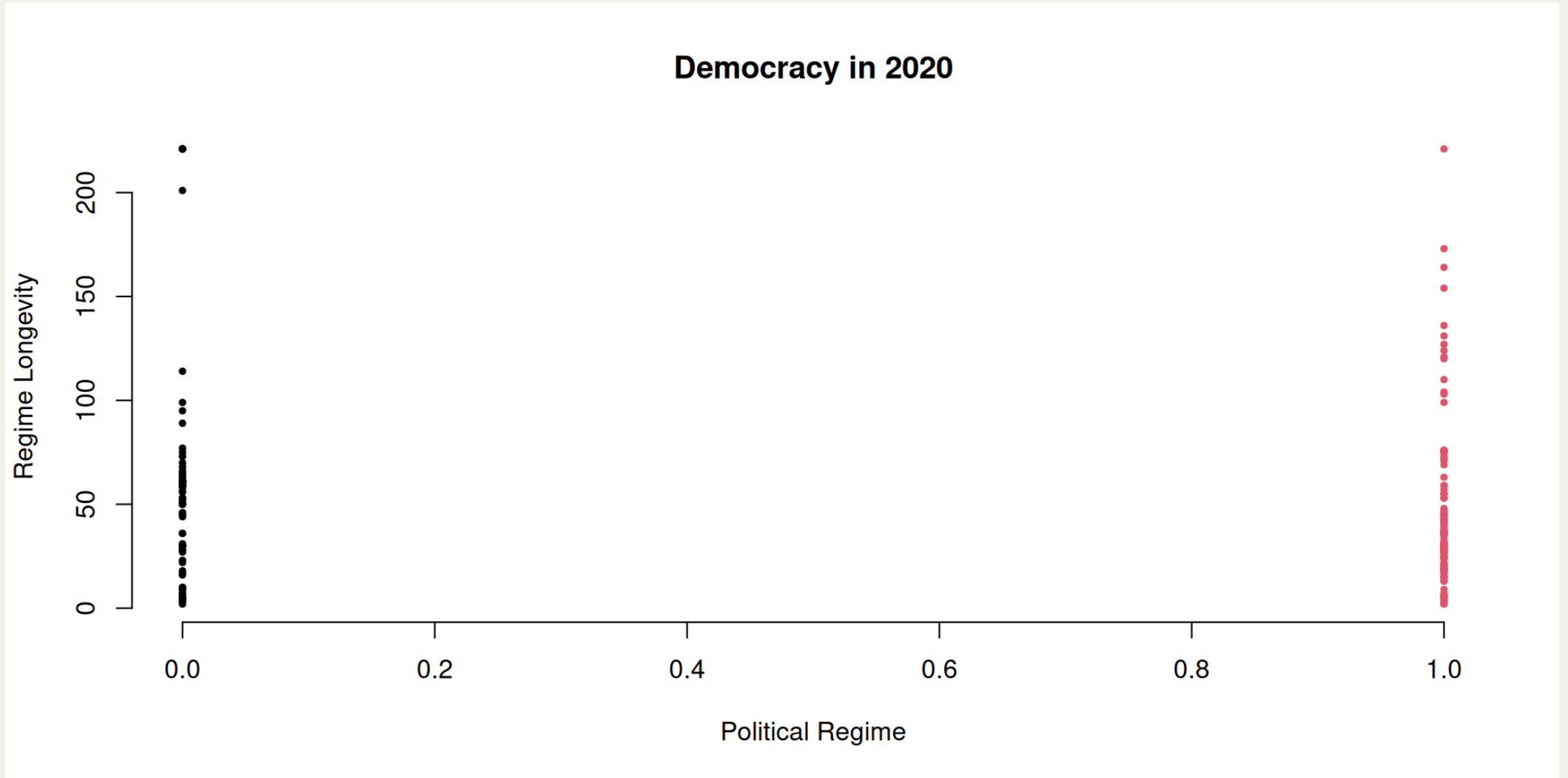
```
1 plot(democracy_gdp_2020$democracy, democracy_gdp_2020$democracy_duration)
```



Prettifying Scatter Plot

```
1 plot(  
2   x = democracy_gdp_2020$democracy,  
3   y = democracy_gdp_2020$democracy_duration,  
4   xlab = "Political Regime",  
5   ylab = "Regime Longevity",  
6   main = "Democracy in 2020",  
7   pch = 19, # Solid points  
8   cex = 0.5, # Smaller points  
9   bty = "n" # Remove surrounding box  
10 )
```

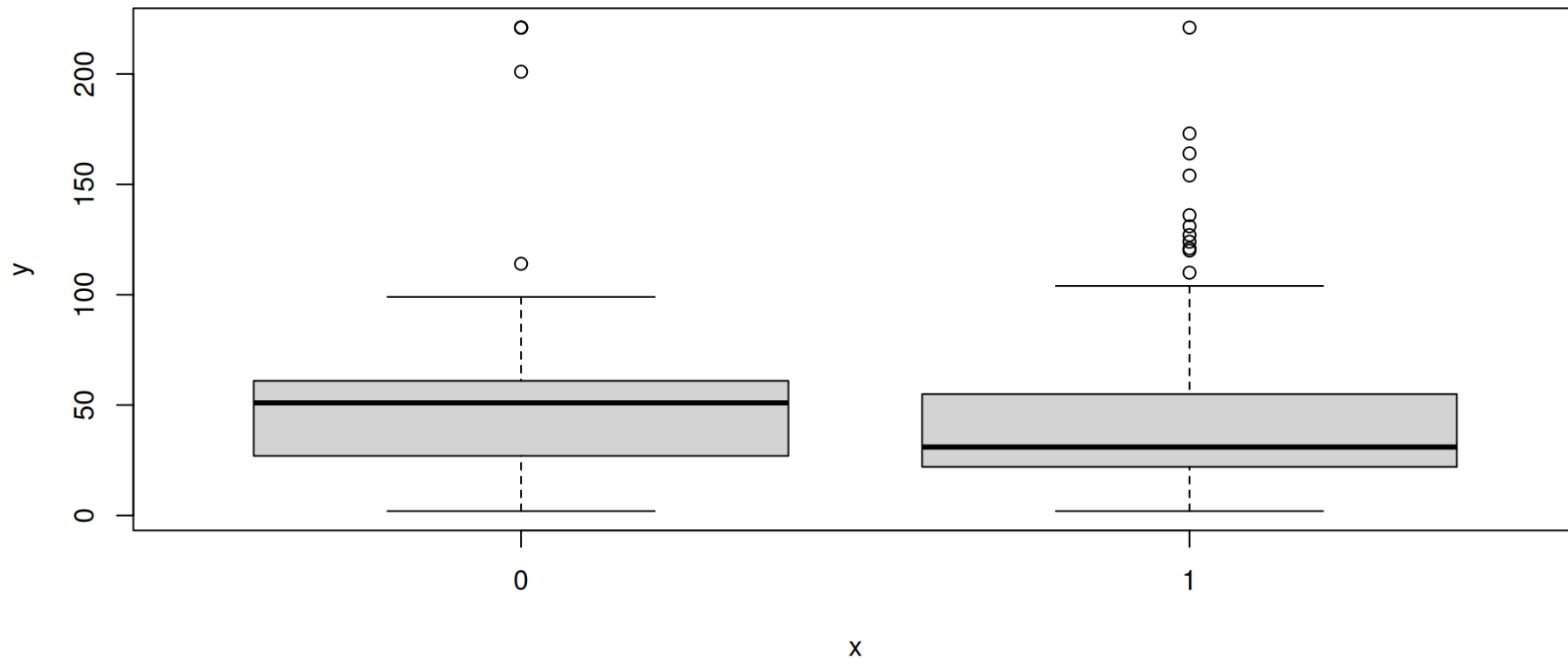
Prettifying Scatter Plot



Box Plot Example

```
1 # Convert democracy indicator into factor
2 democracy_gdp_2020$democracy <- factor(democracy_gdp_2020$democracy)

1 plot(democracy_gdp_2020$democracy, democracy_gdp_2020$democracy_duration)
```



Comparing Pizzas



Fermat's Library
@fermatlibrary



To celebrate National Pizza Day we decided to revisit this fun fact.

Here's a useful counterintuitive fact: one 18 inch pizza has more 'pizza' than two 12 inch pizzas



$$\text{Area} = \pi(18/2)^2 = 254 \text{ in}^2$$



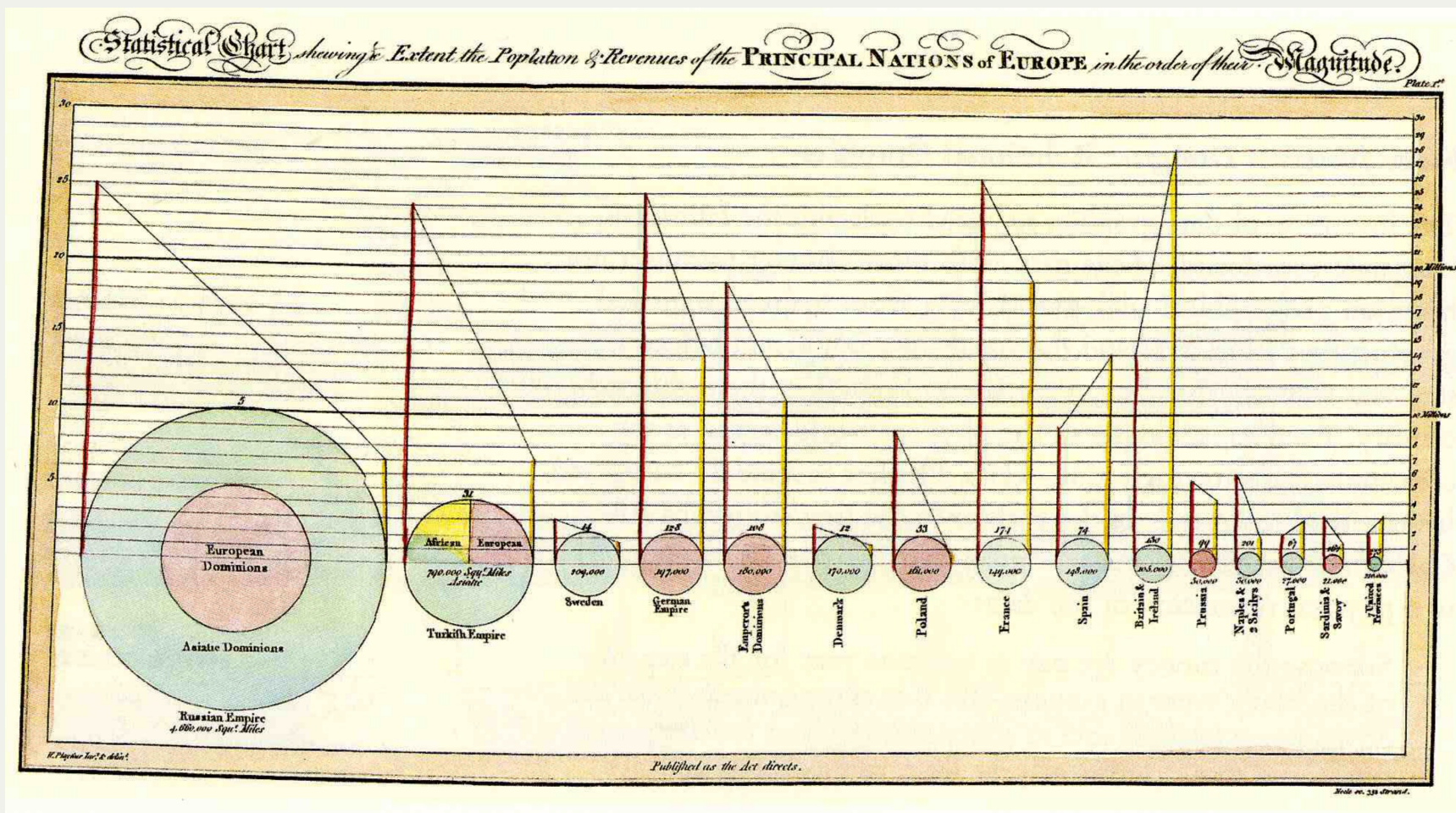
$$\text{Area} = 2\pi(12/2)^2 = 226 \text{ in}^2$$

2:35 pm · 10 Feb 2022 · Buffer

X (Twitter)

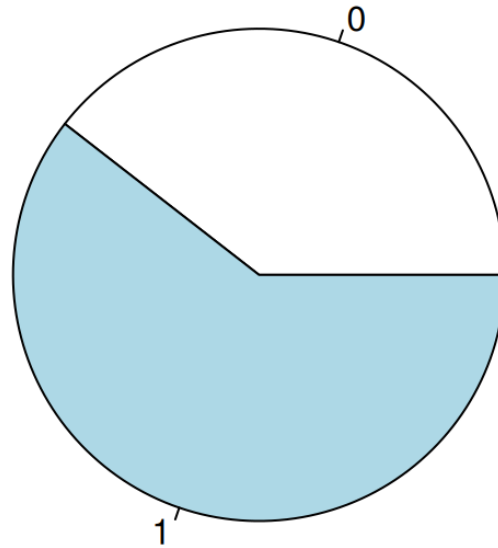
Pie Charts

- While broadly popular, pie charts are generally not a good way of presenting information.



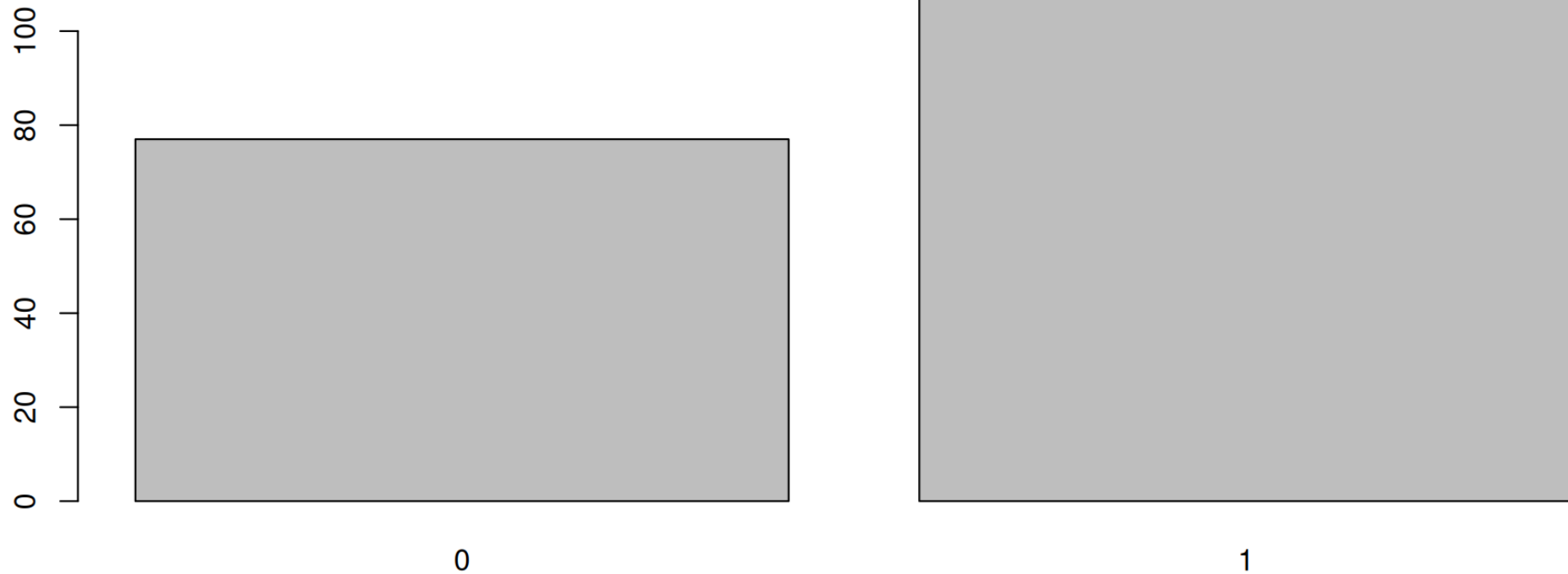
Pie Charts in R

```
1 pie(table(democracy_gdp_2020$democracy))
```



Use Bar Plot Instead

```
1 barplot(table(democracy_gdp_2020$democracy))
```



Beyond Base R

- While extremely powerful, base R plotting facilities are not as flexible as some external libraries.
- `ggplot2` is particularly popular for data visualisations.
- It is based on the Grammar of Graphics plotting approach:
 - Graphs are broken into multiple layers
 - Layers can be recycled across multiple plots
- As an external library `ggplot2` needs to be installed and loaded before using:

```
1 install.packages("ggplot2")  
2 library("ggplot2")
```

Structure of ggplot calls

- Creation of ggplot objects in plotline has the following structure:

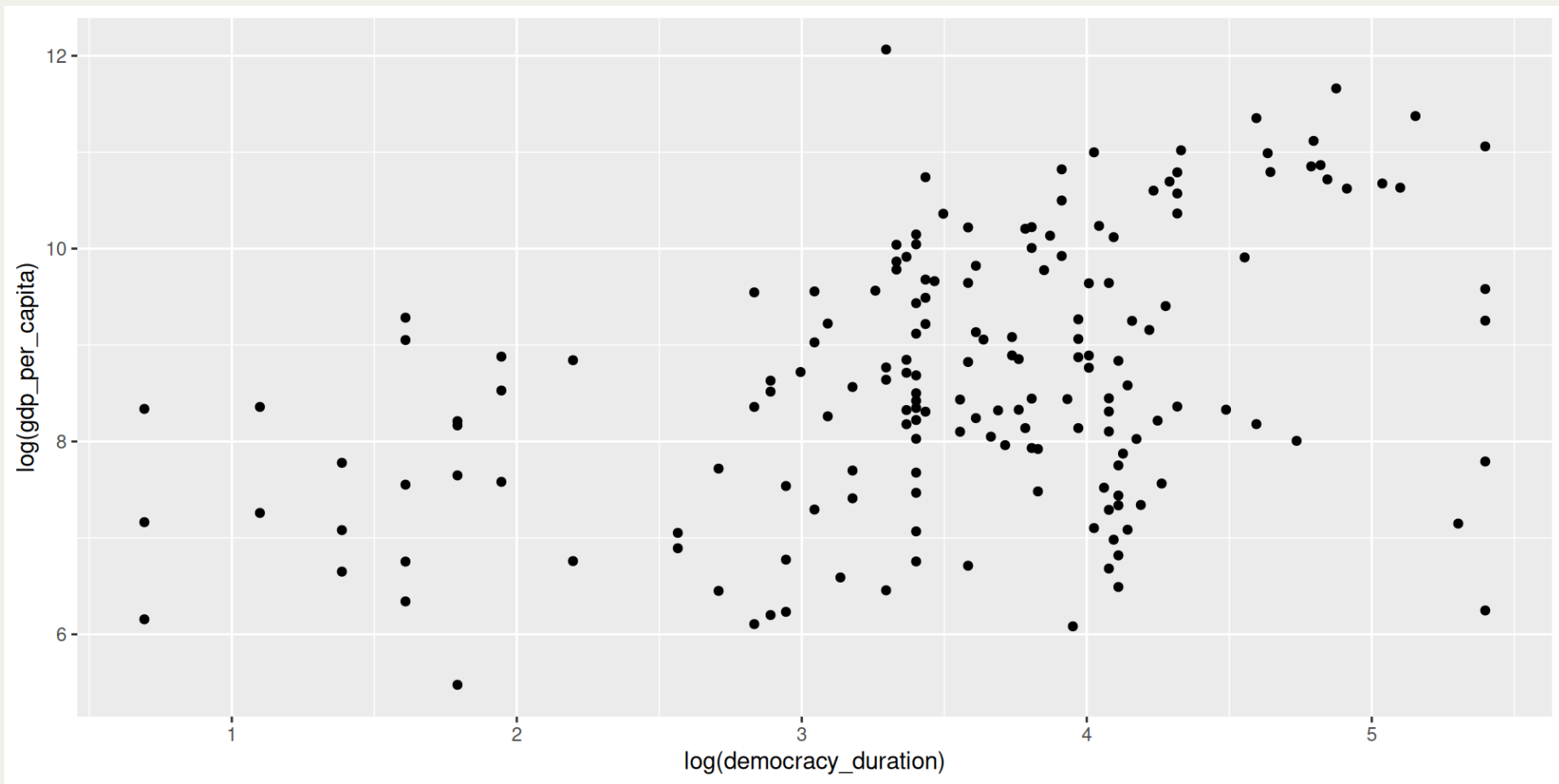
```
ggplot(data = <data>) +  
  <geom_function>(mapping = aes(<mappings>))
```

- If the mappings are re-used across geometric objects (e.g. scatter plot and line):

```
ggplot(data = <data>, mapping = aes(<mappings>)) +  
  <geom_function>() +  
  <geom_function>()
```

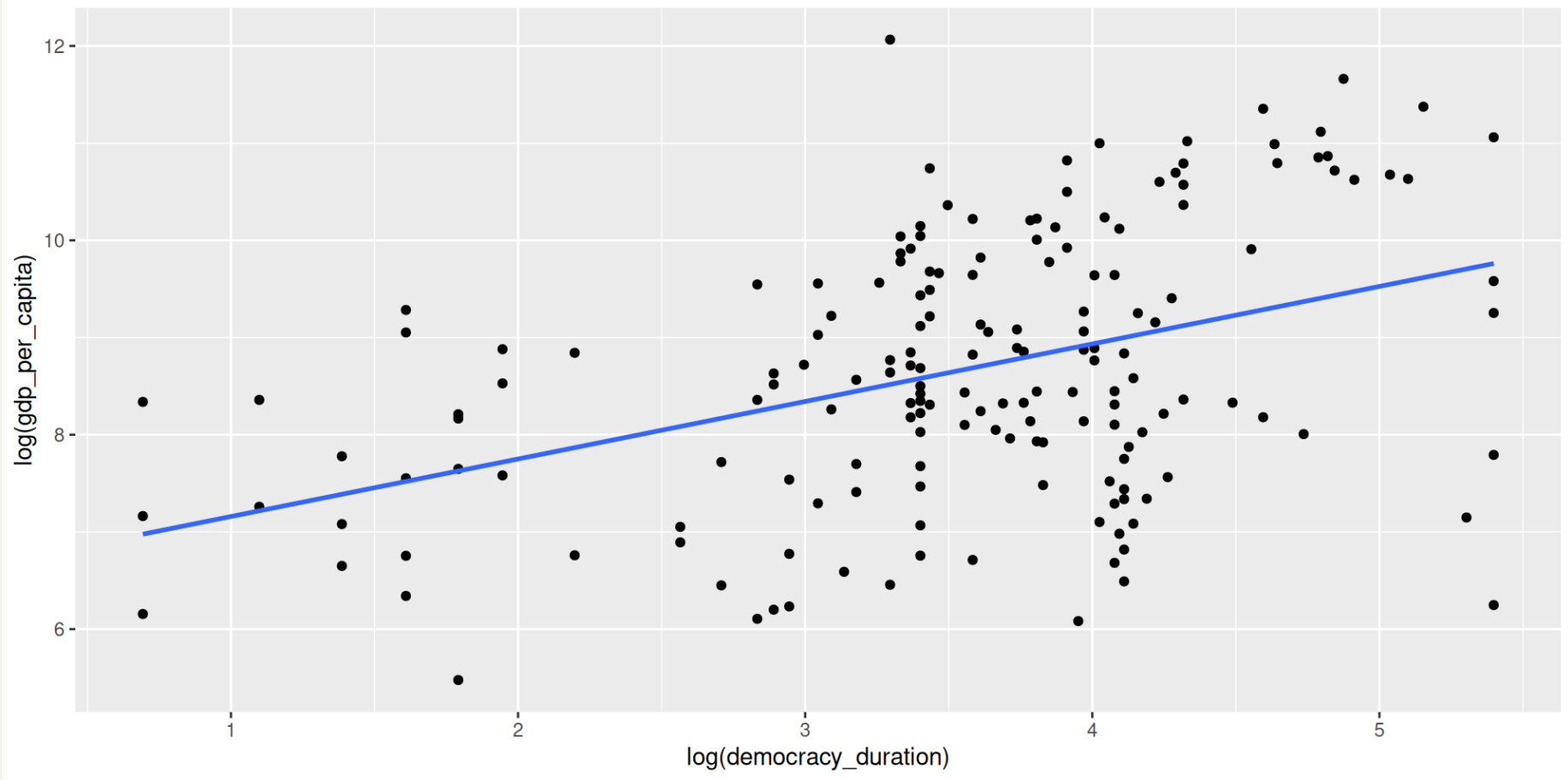
ggplot Example

```
1 democracy_plot <-  
2   ggplot(democracy_gdp_2020, aes(x = log(democracy_duration), y = log(gdp_p  
3   geom_point()  
4   democracy_plot
```



Adding Layers

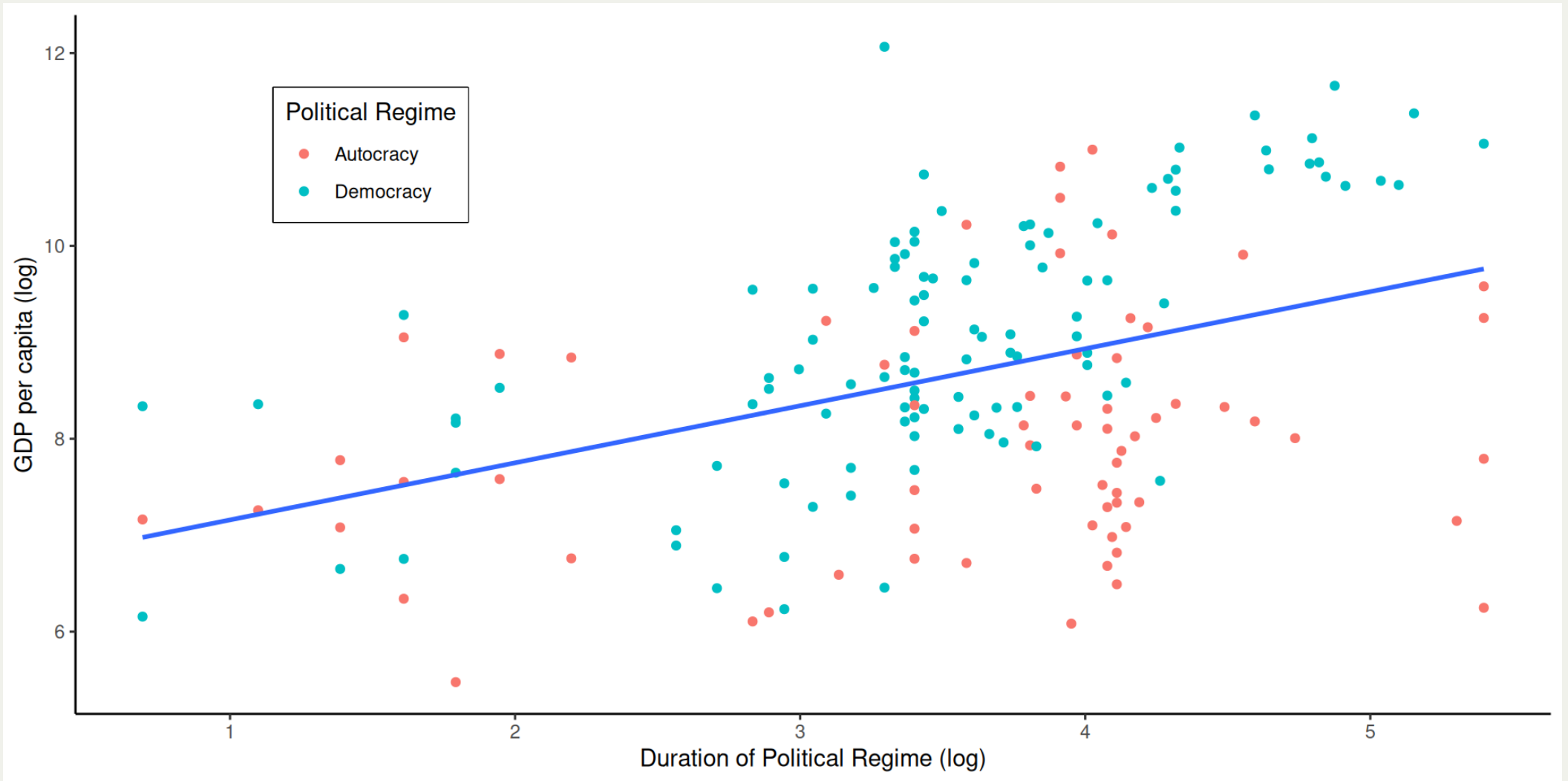
```
1 # Re-use previously saved plot
2 democracy_plot +
3   # Add a layer with linear regression line
4   geom_smooth(method = lm, se = FALSE)
```



Prettifying ggplot

```
1  ggplot(  
2    democracy_gdp_2020,  
3    aes(x = log(democracy_duration), y = log(gdp_per_capita))  
4  ) +  
5    geom_point(  
6      aes(colour = factor(democracy, labels = c("Autocracy", "Democracy")))  
7    ) +  
8    geom_smooth(method = lm, se = FALSE) +  
9    labs(  
10     x = "Duration of Political Regime (log)",  
11     y = "GDP per capita (log)",  
12     colour = "Political Regime"  
13   ) +  
14   theme_classic() +  
15   theme(  
16     legend.box.background = element_rect(size = 0.5),  
17     legend.position = c(.2, .8)  
18   )
```

Prettifying ggplot



Available Geometric Objects

Method	Description
<code>geom_bar()</code> , <code>geom_col()</code>	Bar charts
<code>geom_boxplot()</code>	Box and whisker plot
<code>geom_histogram()</code>	Histogram
<code>geom_point()</code>	Scatterplot
<code>geom_line()</code> , <code>geom_path()</code>	Lines
<code>geom_map()</code>	Geographic areas
<code>geom_smooth()</code>	Smoothed conditional means
<code>geom_violin()</code>	Violin plots

Next

- Tutorial:
 - Correlation
- Next week:
 - Reading week
- After reading week:
 - RQ Presentations