Unit 3 - Graphs

Flinta* R-Tutorium

WU Wien

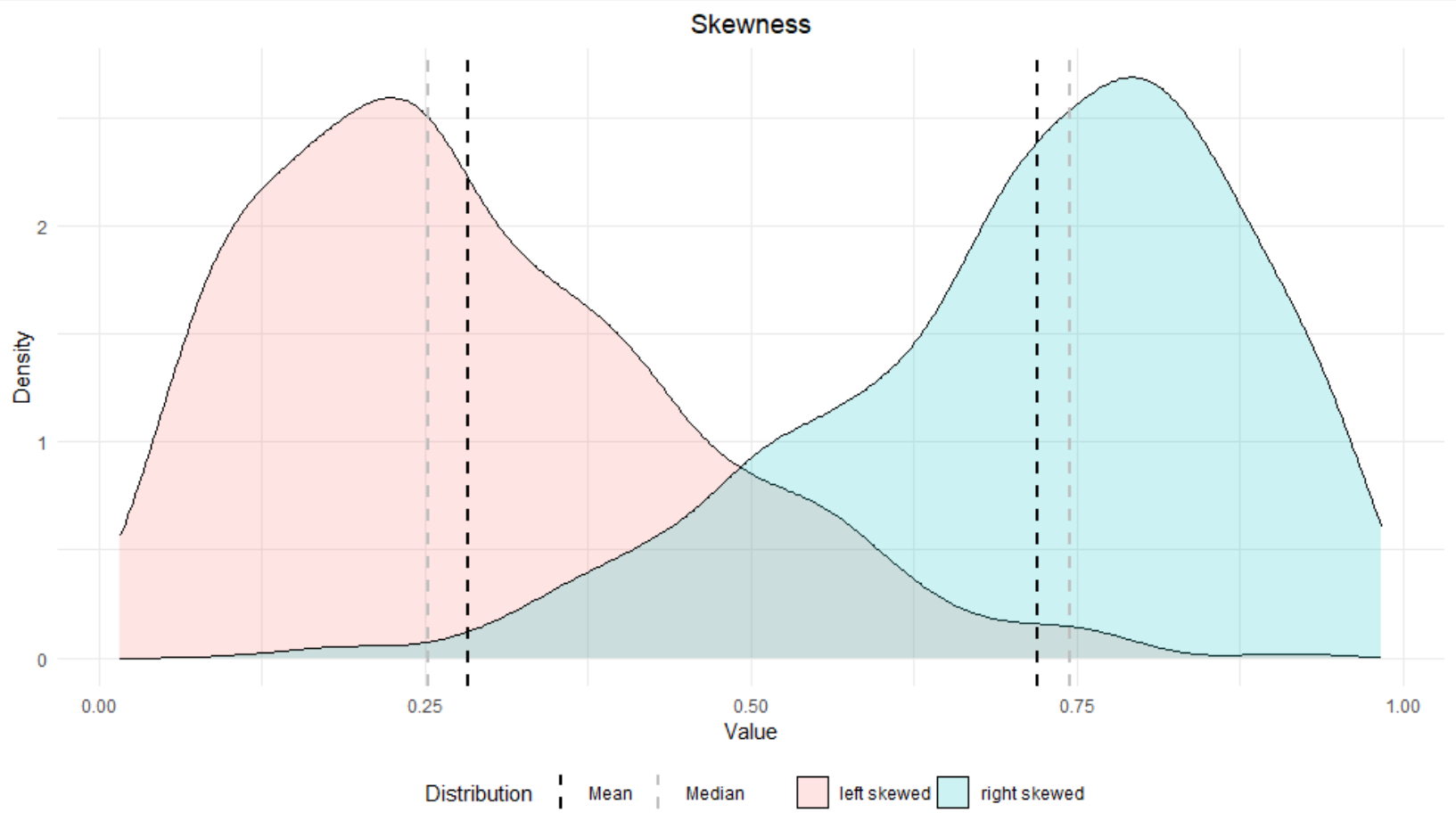
08.04.2024

Numbers are nice - Graphs are better

Numbers

1. Mean = 0.70
2. Median = 0.73
3. Standard deviation: 0.1627
4. Skeweness = -0.64

Numbers are nice - Graphs are better



How to do it in R

1. *ggplot* is the most important command.
2. It is very versatile
3. I will guide you through the command, but it is impossible to cover all the possibilities of ggplot today.
4. A really useful overview can be found at ggplot's website `ttps://ggplot2.tidyverse.org/`
5. But don't worry ChatGPT knows the ggplot also very well

Syntax ggplot

ggplot(dataset name, aes(x-axis, y-axis)) +
geom_point()

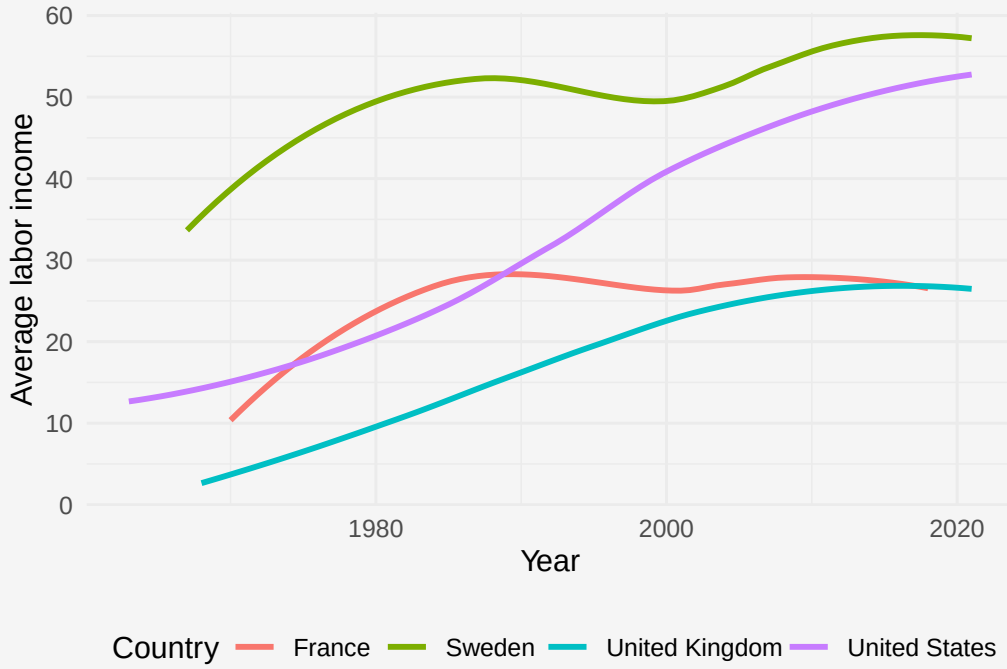
1. ggplot = indicates that we want to use the ggplot command
2. aes = which variables are shown.
3. geom_point = how data is displayed in this case by points (but you can also create lines, bars..)

Here is the link[1] to a cheat-sheet for a quick overview of different geom_ options.

[1] https://github.com/rstudio/cheatsheets/blob/main/data-visualization.pdf

ggplot example

Using LIS (2023) data to depict the trend of the average labor income of the top 10% capital earners




```
ggplot(data, aes(year, avgL, group = country, color = country)) +
geom_smooth() +
labs(x = "Year", y = "Average labor income percentile") +
theme_minimal() +
theme(legend.position = "bottom")
```

Further links

1. https://web.stanford.edu/~lstell/ggplot2Intro.pdf
2. https://r-statistics.co/Complete-Ggplot2-Tutorial-Part1-With-R-Code.html?utm_content=cmp-true

Aesthetics:

- http://www.cookbook-r.com/Graphs/Colors_(ggplot2)/
- https://github.com/karthik/wesanderson