Using `ggplot2` for Data Visualisation

R. Riegler and M. Hisarciklilar

Introduction

- **Part 1: Background**
  - What is Grammar of Graphics (GoG) and why do we need it?
  - What is `ggplot2`?
- **Part 2: Hands-on activities**
  - `ggplot2` basics
  - Student exercise
  - More dimensional graphics
  - Animations using `gganimate`

Grammar of Graphics (GoG)

- Leland Wilkinson (1999) introduced the *Grammar of Graphics* (GoG)
- GoG is a structured approach of turning data into visualizations by following a specific set of steps.
- Each step plays a unique role, like preparing the data, choosing how to display it, and adjusting its appearance.
- Using these steps, we can create a wide variety of meaningful charts and graphs.
- How can it be operationalised?

ggplot2

- Hardley Wickham (2005)'s `ggplot2` package follows the principles of GoG, and applies the idea of a *layered* graphics generation using R.
- Rationale is similar to building a house:
  - **Foundation**: We select a dataset that we want to use for plotting. We decide what variable we put on the axes, and set some global *aesthetics* options that will apply for all the layers we use.
  - **Building the house**: To actually produce the chart, we have to add geometric objects (geoms), such as bars, lines, points, etc. We could add another data layer to the graph with its own aesthetics, which is similar to adding a conservatory with its own style to your house.
  - **Decorating**: To make sure that the graph looks perfect, we can modify the scales, labels and legend of the graph. We can also apply themes and faceting.

Some good practice...

- Open R and R Studio.
- Create a folder, which will be used as our working directory.
- Set your working directory in R.
- Open an R-Script.
- We are ready for the first exercise...

Introduction to ggplot2 using mpg data

- `cty`, **hwy**: miles per gallon (mpg) for city and highway driving.
- **displ**: engine displacement in litres.
- `drv`: drivetrain; front wheel (f), rear wheel (r) or four wheel (4).
- `model`: model of car.
- **class**: categorical variable describing the "type" of car: two seater, SUV, compact, etc.

Student Task I - Regional Data

Import the *region_data.csv* dataset. This ONS dataset contains aggregated regional data on gross disposable household income per capita for the period 1997 to 2022. Data is measured in GBP at current prices.

Complete the following tasks:

1. Look at the data

Student Task I - Regional Data

2. Create a plot using the ggplot command. Add *year* to the x axis, *GDHI_pc* to the y-axis and the `colour = Region_name` option to the mapping aesthetics to illustrate regional differences. Select a geometric object from the list below to display the data. Which one is most suitable for?

- `geom_area`
- `geom_line`
- `geom_point`

4. Use faceting to create a plot for each region. Surpress the legend by adding the layer `theme(legend.position="none")`.
5. In which regions did we find the largest increase in gross disposable household income per capita?

Themes

- Existing themes can help you to create plots that follow a specific style.
- E.g. you can create charts that looks like charts used by the BBC.
- Load ***region_data_sel.csv***

Bubble Charts: Income and Life-Expectancy

- We are now creating a more advanced plot to display the relationship between income and life-expectancy
- Load bubble.chart.data.csv

Animations: Income and Life Expectancy

- For our last exercise, we want to introduce animations.
- Load animate.csv

Useful resources

- Different shape codes: https://ggplot2.tidyverse.org/articles/ggplot2-specs.html#sec:shape-spec
- Colour schemes: https://www.w3schools.com/colors/colors_picker.asp

ggplot2 References

- Davies, T. M. (2016). The book of R: a first course in programming and statistics. No Starch Press.
- Keyes, D. (2024). R for the rest of us: a statistics free introduction. No Starch Press. https://rfortherestofus.com/book
- Wickham, H., Navarro, D., and Pedersen, TL (2025) ggplot2: Elegant Graphics for Data Analysis (3e). https://ggplot2-book.org/
- Data.Europa.EU (2023). Data Visualisation - A Comprehensive Guide to Unlocking Your Data's Potential https://data.europa.eu/apps/data-visualisation-guide/