Introduction & overview

Theoretical and Empirical Research Methodology,
Implementation Lab 1

Prof. Dr. Claudius Gräbner-Radkowitsch
Europa-University Flensburg, Department of Pluralist Economics
www.claudius-graebner.com | @ClaudiusGraebner | claudius@claudius-graebner.com

Europa-Universität Flensburg

Europa-Universität Flensburg
International Institute of Management and Economic Education
Department of Pluralist Economics

Part I:

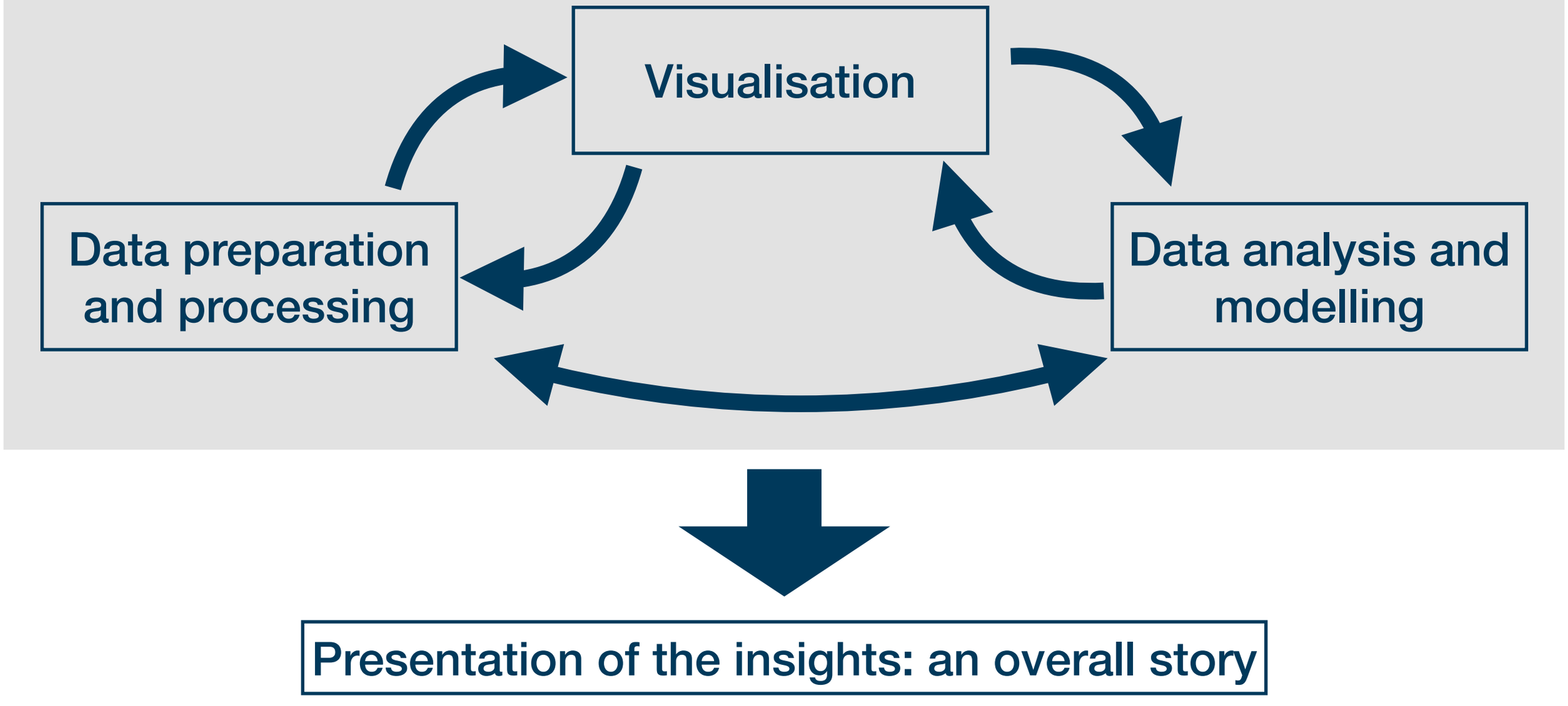
Organization & outlook

Note: my slides for this course are meant as a “script on slides”

The goal of the implementation lab

Learn how to prepare, analyse, and present quantitative data using **R**

Implement all tools covered in the main lecture practically



Why R?

- R allows you to conduct **all steps of this data science pipeline** within one consistent framework in a transparent and reproducible manner

- R is free, OS-independent and **open source**
  → inclusive, transparent, and vibrant tool

> “The days of commercial statistical languages and packages such as SAS, Stata and SPSS are over”
>
> Paul Jansen, CEO of Tiobe Software

- For statistical analysis, R is among the **most widely used** and demanded programming languages

- R is demanded in almost **every industry**

- Learning R makes it **easier to learn other** widely used programming languages

- There is a great and friendly R **Community**

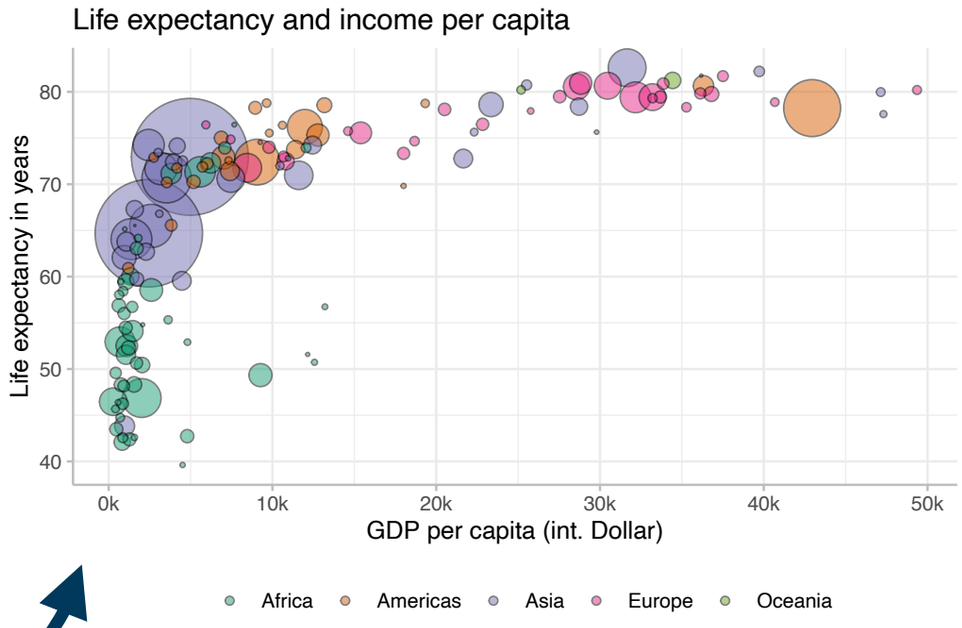
#	RedMonk	TIOBE	PYPL
1	JavaScript	Python	Python
2	Python	C	Java
3	Java	Java	JavaScript
4	PHP	C++	C/C++
5	C#	C#	C#
6	C++	Visual Basic	PHP
7	CSS	JavaScript	R
8	TypeScript	PHP	Objective C
9	Ruby	Assembly	Swift
10	C	SQL	TypeScript
11	Swift	Go	Matlab
12	R	Swift	Kotlin
13	Objective C	R	Go
14	Shell	Matlab	Ruby
15	Scala	Delphi	VBA

What you will be able to do

- **Read** in data sets from various sources
- **Prepare** ‘messy’ data and produce ‘tidy’ data
- Create illustrative **visualisations** on a publication-ready level

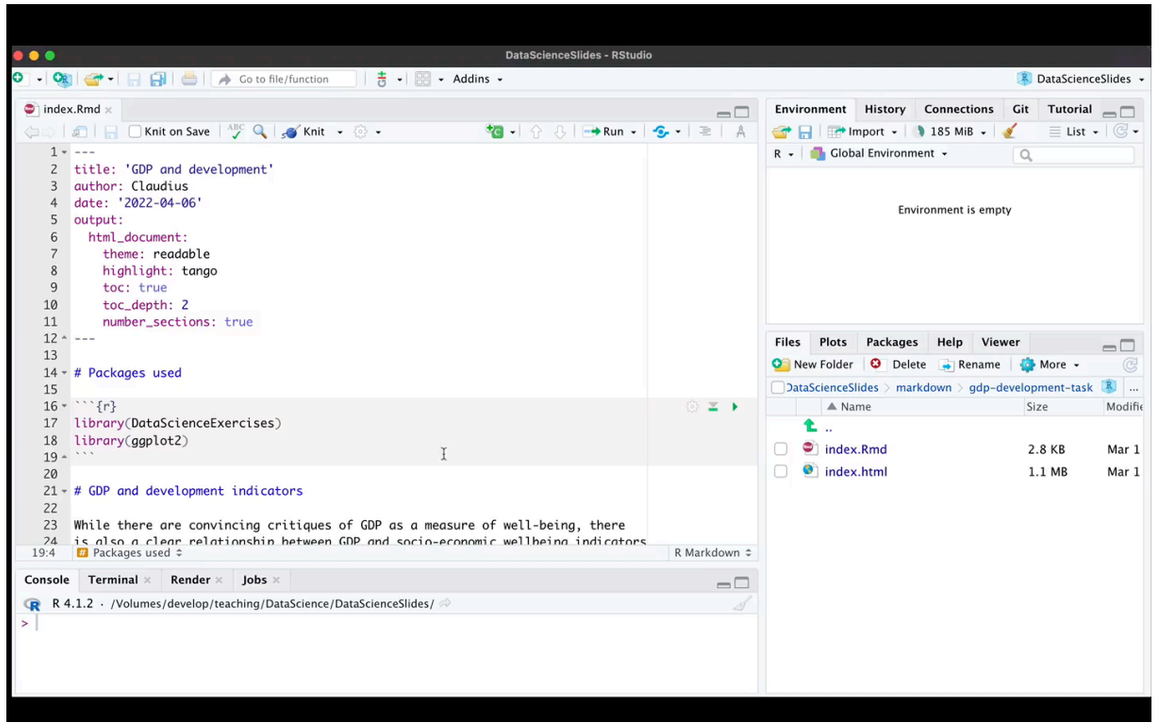
```
country,1952,1957,1962,1967,1972,1977,1982,1987,1992,1997,2002,2007
Afghanistan,Asia|28.801|8425333|779.4453145,Asia|30.332|9240934|820
.8530296,Asia|31.997|10267083|853.10071,Asia|34.02|11537966|836
.1971382,Asia|36.088|13079460|739.9811058,Asia|38.438|14880372|786
.11336,Asia|39.854|12881816|978.0114388,Asia|40.822|13867957|852
.3959448,Asia|41.674|16317921|649.3413952,Asia|41.763|22227415|635
.341351,Asia|42.129|25268405|726.7340548,Asia|43.828|31889923|974
.5803384
Albania,Europe|55.23|1282697|1601.056136,Europe|59.28|1476505|1942
.284244,Europe|64.82|1728137|2312.888958,Europe|66.22|1984060|2760
.196931,Europe|67.69|2263554|3313.422188,Europe|68.93|2509048|3533
.00391,Europe|70.42|2780097|3630.880722,Europe|72|3075321|3738
.932735,Europe|71.581|3326498|2497.437901,Europe|72.95|3428038|3193
.054604,Europe|75.651|3508512|4604.211737,Europe|76.423|3600523|5937
```

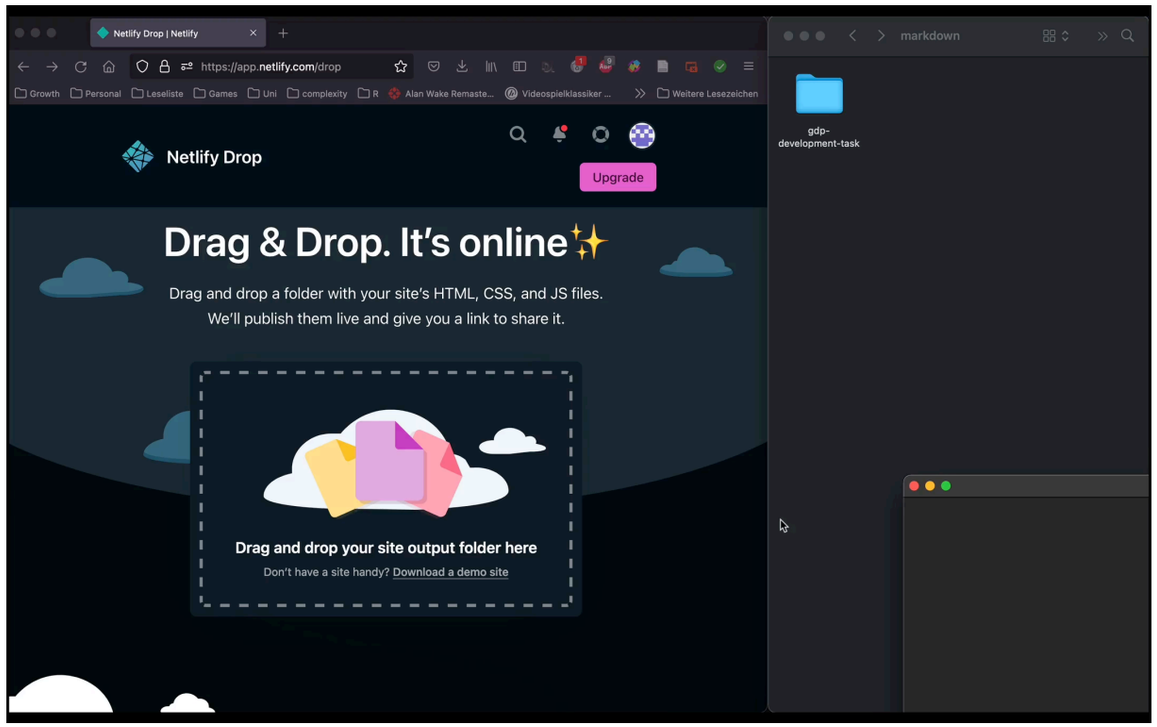
```
# A tibble: 142 × 5
   country       continent lifeExp       pop gdpPercap
   <fct>         <fct>       <dbl>     <int>     <dbl>
 1 China         Asia         73.0 1318683096     4959.
 2 India         Asia         64.7 1110396331     2452.
 3 United States Americas     78.2  301139947    42952.
 4 Indonesia     Asia         70.6  223547000     3541.
 5 Brazil        Americas     72.4  190010647     9066.
 6 Pakistan      Asia         65.5  169270617     2606.
 7 Bangladesh    Asia         64.1  150448339     1391.
 8 Nigeria       Africa       46.9  135031164     2014.
 9 Japan         Asia         82.6  127467972    31656.
10 Mexico        Americas     76.2  108700891    11978.
# … with 132 more rows
```



What you will be able to do

- **Identify hidden patterns** in data and **make predictions** using models
- Write reproducible research **reports** in Quarto
- **Publish** visually appealing reports on the web via Netlify
- You learn tools and methods that will be required in subsequent seminars





But can’t ChatGPT just do this for us…?

- Yes and no 😇
- LLMs such as ChatGPT can be very useful when developing code
- You know a programming language well → LLMs boost your productivity
- You do not know a language well → LLMs will be confusing
- Often return *slightly* misbehaving code → always need to work on the output
- LLMs do not know the most recent developments in the language
- Their code is not always easy to read

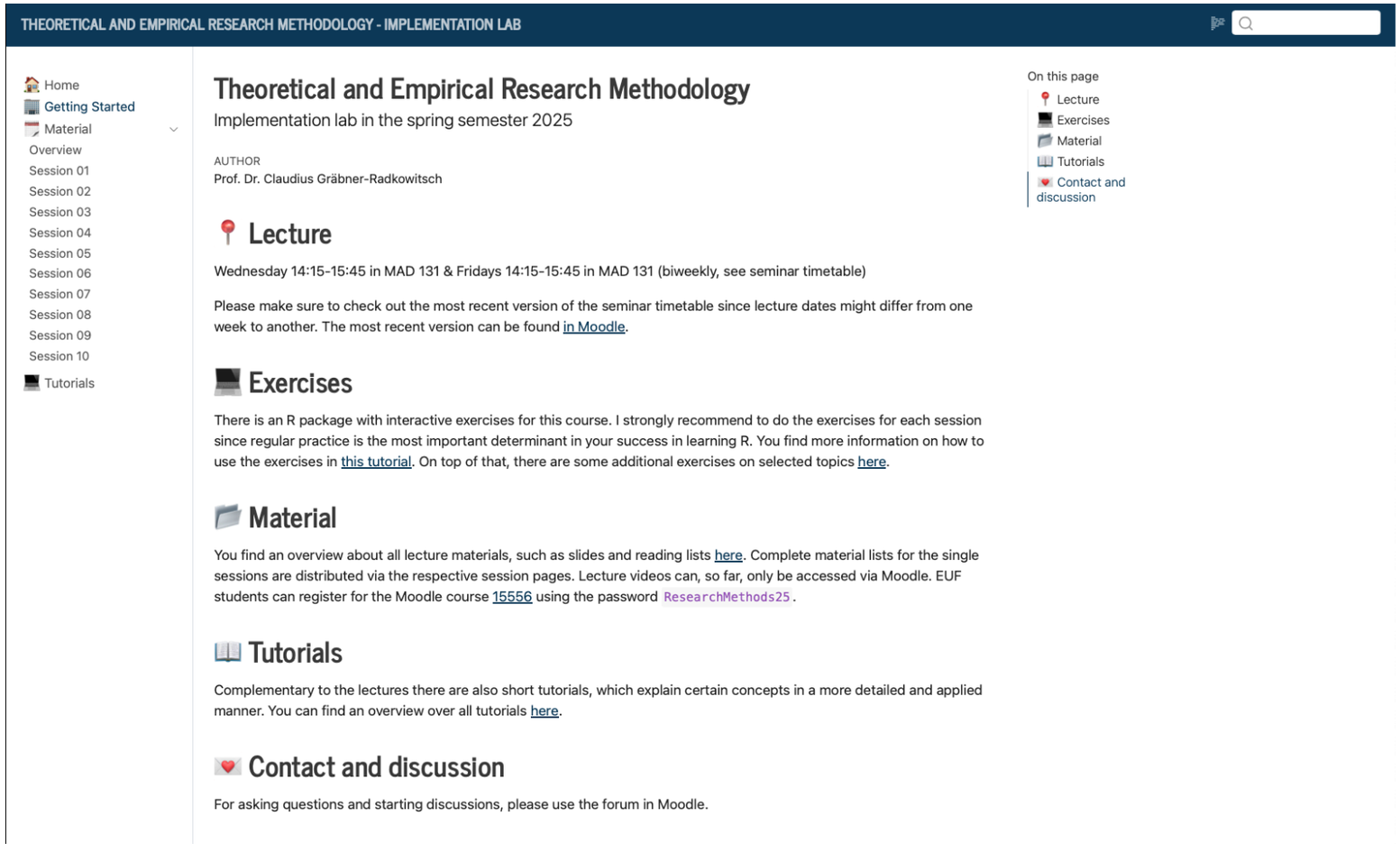
We will learn the basics of using LLMs for coding and you can always use them, including during the exam, but be wary of their limitations 🤓

Organization of the lab sessions

- Each on-site session comprises theory and practice → always bring **laptops**👩🏼‍💻👩🏾‍💻
  - Some topics are presented better by **instructional videos** → watch the videos, do the **exercises**, pose your questions
- Questions should always be posted online in the **Moodle forum**
  - Questions should primarily be **answered by fellow students** → helping one another with problems greatly enhances understanding.
  - The forum ensures that answers to questions are (i) recorded and (ii) available to everybody
  - Particularly intriguing questions can be discussed at the beginning of a session

A short tour through the website

- Two main areas:
  - Material
  - Tutorials
- Don’t forget about the exercises
- Note: always reload the website to make sure you are on the most recent version



Some remarks on our ‘learning agreement’

- Why do I expect regular activities from you?
  - Learning a programming language is a **consecutive activity**: you miss basics in the beginning → you’ll quickly become frustrated and get lost
  - This is a demanding course: catching up later on what you missed earlier will be difficult
  - Learning a programming language works mainly through practice and **doing** → practical exercises have a *huge* benefit
  - Learning a programming language is *difficult* and at times *frustrating* → we need an amicable environment and must support each other
  - Few things have a bigger learning effect than helping others with their problems

Learning a programming language can be a lot of fun and really brings you forward – if we do this together as a team💪🏽

Open questions?

Part II:

Installing R and R Studio

R and R-Studio

- R is a programming language
  - It is a language that allows you to issue commands to your computer:

```
> fib_n(4)
```

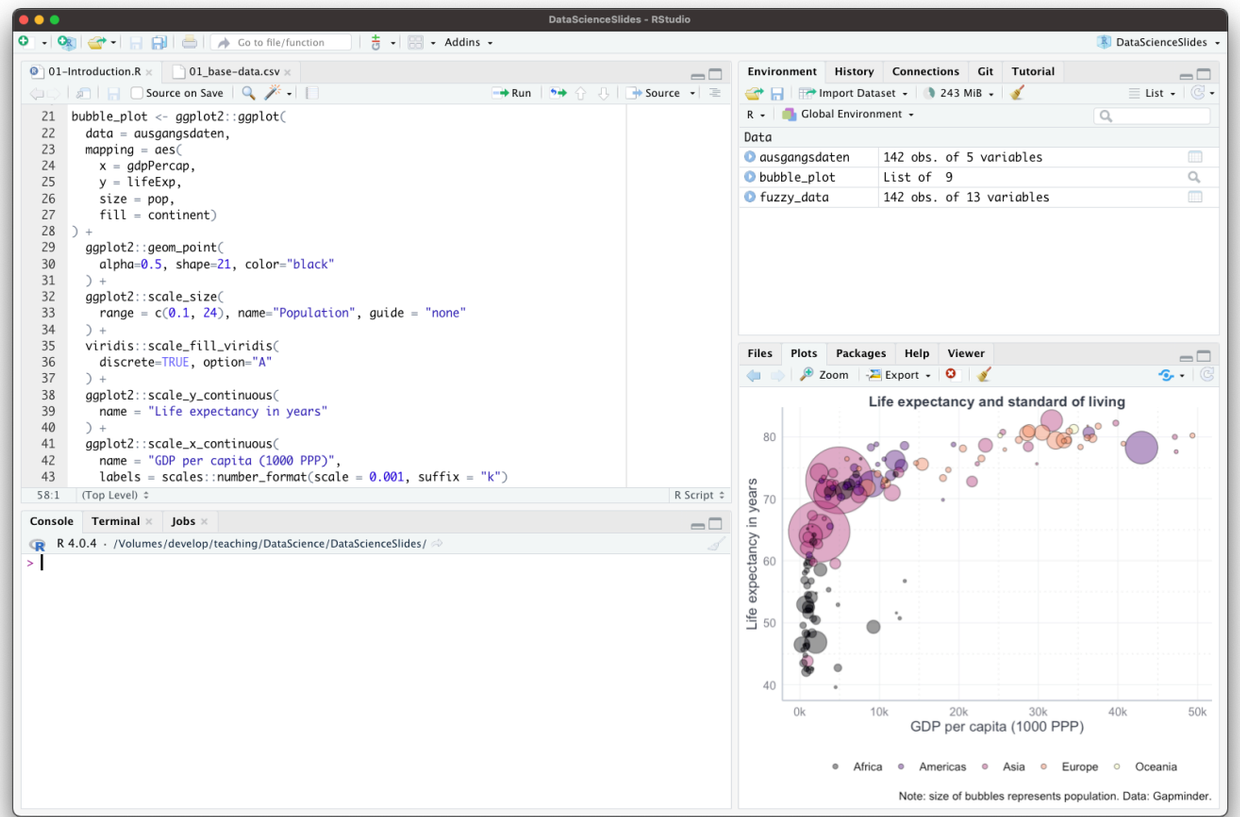


```
8B542408 83FA0077 06B80000 0000C383
FA027706 B8010000 00C353BB 01000000
B9010000 008D0419 83FA0376 078BD989
C14AEBF1 5BC3
```



```
[1] 3
```

- R-Studio is an integrated development environment
  - Basically a fancy text editor with additional features that make programming easy



R and R-Studio

- R is a programming language
- R-Studio is an integrated development environment

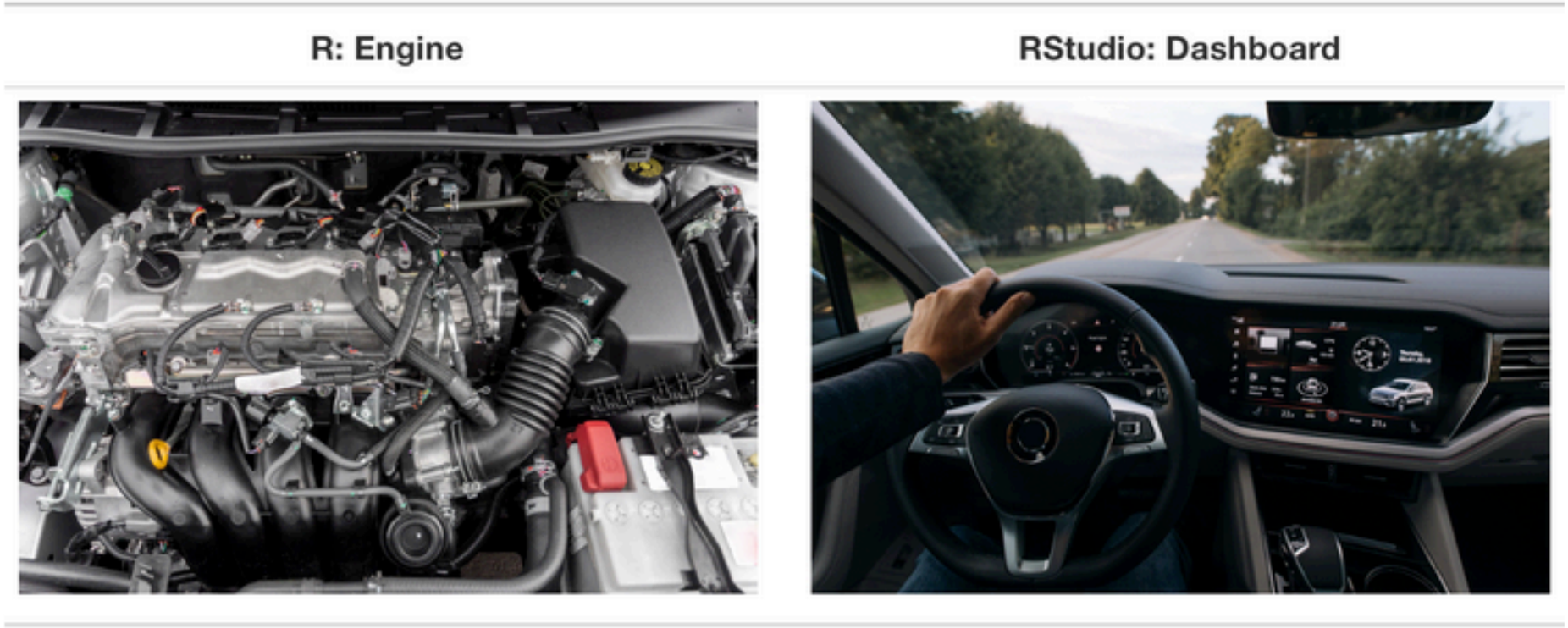


Figure: Ismay & Kim (2022)

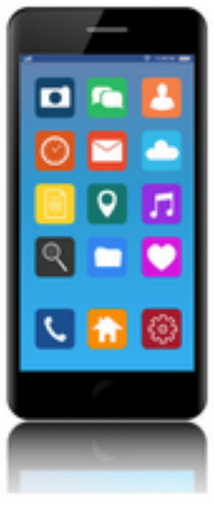
- You need to install R first, then you can install R Studio
- After that, you basically only use R Studio → it calls R whenever necessary

R and R packages

- If you install R, you can issue a lot of commands that your computer immediately understands
- However, there might be some routines that R “doesn’t understand”
- You might “teach” R this by defining, for instance, certain functions that perform these operations
- You might then even “save” these functions and pass it on to others, so that they can use them as well
- This is the idea of **R packages**: a collection of variables and functions written by others that you can install on your computer and use them
- Once an R package is installed, you can use all functions and variables defined by the creator of the package

R and R packages

- Again, Ismay & Kim (2022) have a nice analogy:

R: A new phone	R Packages: Apps you can download
	

- I wrote a small script that installs all packages that we will use throughout the semester, so we can already resolve all installation issues now

And what about LaTeX?

- In this course we learn how to write nice reports in Quarto
  - Put R code and text into one file → get a webpage in HTML or a nice PDF
- Creating HTML code is easy, creating a PDF is nothing trivial
  - To do this, we need a software called LaTeX → a typesetting system
  - It turns plain text into nice text within a PDF document

```
---
title: "R Markdown and PDF"
author: "Claudius"
output: pdf_document
---

Here I write text. And add code:

```{r}
sqrt(1:5)
```
```

```
\ifLuaTeX
  \usepackage{selnolig}  % disable illegal ligatures
\fi

\title{R Markdown and PDF}
\author{Claudius}
\date{}

\begin{document}
\maketitle

Here I write text. And add code:

\begin{Shaded}
\begin{Highlighting}[]
\FunctionTok{sqrt}\NormalTok{(}\DecValTok{1}\SpecialCharTok{
:}\DecValTok{5}\NormalTok{)}
\end{Highlighting}
\end{Shaded}

\begin{verbatim}
## [1] 1.000000 1.414214 1.732051 2.000000 2.236068
\end{verbatim}
```

This is LaTeX

R Markdown and PDF

Claudius

Here I write text. And add code:

```
sqrt(1:5)

## [1] 1.000000 1.414214 1.732051 2.000000 2.236068
```

Installation procedure

- It is absolutely essential that you install all the necessary software as soon as possible → installation guidelines on the course homepage
- Until next session you should have…
  - …tried to install R, R Studio and Git → follow my tutorials
  - …posted all problems with a screenshot in the Moodle forum

- Be prepared tomorrow, trying to install R just before the session is 🤬
- We need to solve all installation problems until the end of next week
  - Post problems on Moodle, help each other out

Problems with the installation?

1. Check again in the tutorials
2. Post your problems in Moodle
3. Accompany them with screenshots