Lancaster University

Quantitative Research Methods

Matthew Ivory

matthew.ivory@Lancaster.ac.uk

1. Introduction to the course

- Session 1: Introduction to quantitative research methods using R
- Session 2: Data management and data wrangling
- Session 3: Exploratory data analysis
- Session 4: Data visualization
- **Session 5: Live Coding Walkthrough**
- Session 6: Probability and distributions
- Session 7: Tests for discrete variables: Analysing contingency tables
- Session 8: Correlations and t-tests
- Session 9: ANOVA and linear regression
- Session 10: Multiple regression, introduction to generalised linear regression

Session Outline

- 1. R Projects, RMarkdown, and Environment Setup
- 2. Live Data Wrangling Walkthrough

1. Projects, RMarkdown, and Environment

- Workflow tools for data science in R
- This is how I structure and share my analyses
- They aren't essential to use *right now*, but they will improve your workflow and I strongly encourage you to use them

- The follow slides will introduce these ideas because I use them in the live coding example, you can follow along in the slides, but they will be repeated in the worksheet

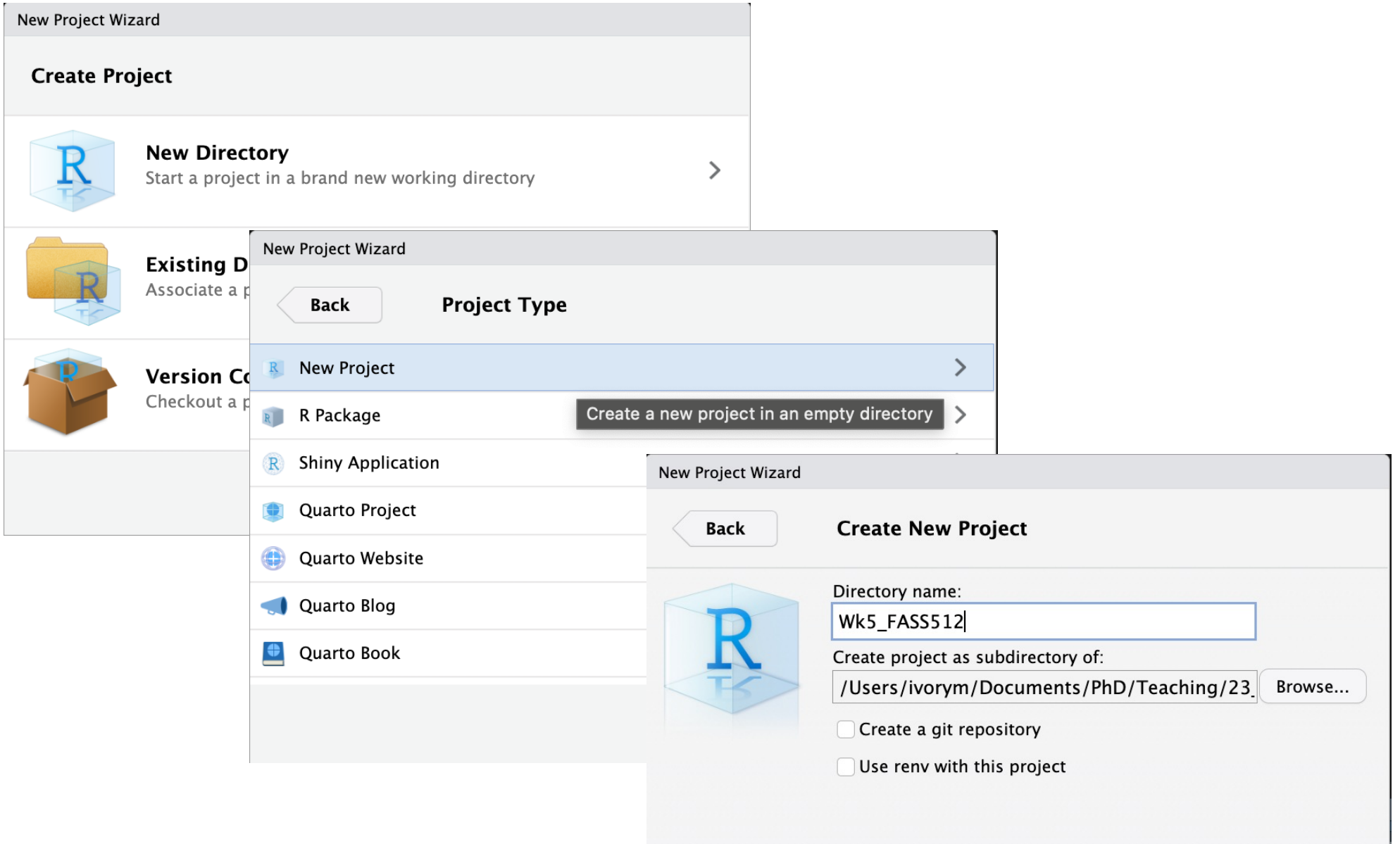
RStudio Projects

- What do you do when you are working with multiple research projects at once?
- What if you need to pause an analysis, wait a few weeks, and then come back?
- How do I keep all my R files together and not have to manually set up my environment each time with setwd(), etc.
- How can I make sure I can share all my analysis files needed?

Working Directory

- Directory = folder
- Your working directory is where R is looking for files
- You can check your working directory by using the command getwd()

RStudio Projects



RStudio Projects

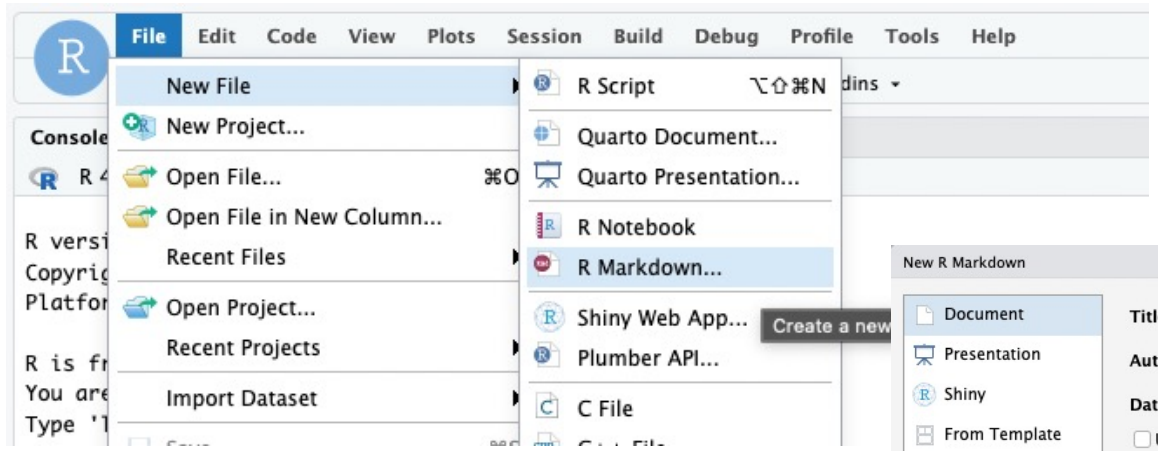
- Once this process is complete, you'll get a new RStudio project just for this book. Check that the "home" of your project is the current working directory using getwd()
- This means R will be looking for files/folders in this folder and below
- It 'centralises' your R scripts and files into one place
- It offers the use of relative, not absolute paths

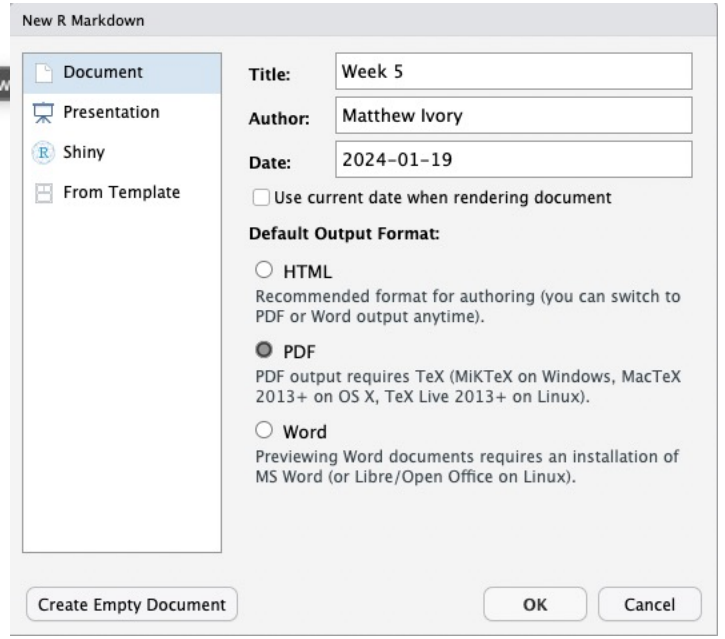
RStudio Projects

- A way to:
  - handle multiple research projects
  - keep all your necessary files together
  - Share your work more easily

RMarkdown

- Like scripts, but ***more***
- R Markdown files are designed to be used for:
  1. For communicating conclusions not the code
  2. For collaborating and sharing both your conclusions, and how you reached them (i.e. the code).
  3. an environment in which we *do* data science, as a notebook where you can record your code, but also what you were thinking, and your research choices.
     - We can do something similar with code comments, but this is just ***more***
- Combines prose with chunks of code
- Can then be compiled/knitted/rendered into a PDF, word document or a webpage






```
1  ---
2  title: "Week 5"
3  author: "Matthew Ivory"
4  date: "2024-01-19"
5  output: pdf_document
6  ---
7
8  ```{r setup, include=FALSE}
9  knitr::opts_chunk$set(echo = TRUE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see
   <http://rmarkdown.rstudio.com>.
15
```

RMarkdown

Lancaster University

- Contains three important types of content
  - YAML (surrounded by ---)
  - Chunks of R code (surrounded by ```)
  - Regular text
- You can mix both text and code
- Code is run directly inside the document below each section
- Helps keep related code together
- Shows plots inside the document instead of the bottom right
- To produce a complete report containing all text, code, and results, click “Knit” or press Cmd/Ctrl + Shift + K.

2. Data Wrangling Walkthrough

- Data taken from Kaggle: https://www.kaggle.com/datasets/kanchana1990/swiss-national-crime-dataset-2008-2022
- Some minor cleaning for the purpose of this task, only looking at 2022 data
- Link to the data is in the worksheet
- Feel free to code along, make notes, ask clarifying questions throughout

2. Data Wrangling Walkthrough

- To RStudio we go…

Questions?

- I will be walking around while you work through the worksheet