1. **Setting up R and Rstudio**

How to find and install R and Rstudio

How to install packages

How to load and check for packages

How to set up Rstudio projects

R Project Management tips

(Advanced): Advanced project management and archiving using Rstudio and GitHub

1. **R for data management**

R as a programming language: Introduction to ‘Base R’

‘Wickhamese R’: How tidyr, dplyr, and related packages make R easier to work with

Code piping and chaining using %>%

Effects and benefits of the tbl\_df data class

1. **Loading data saved as a text file using readr**

Loading a csv file

Loading a tab delimited file

Understanding the arguments of readr functions

Skipping lines

Renaming variables

Specifying variable types

Specifying missing data indicators

Differences between read\_csv and read.csv

**Loading data saved in other file formats**

Loading excel files using readxl

Understanding the structure and arguments of readxl functions

Loading an excel file

Reasons for preferring data in text formats

How to save data in text format within Excel

Loading spss and stata files using haven

Arguments in haven functions

Examples of usage

1. **Initial data tidying using stringr**

R data types: character, factor and numeric

Type conversion issues if data are not tidied

Simple string manipulation: Removing commas and spaces

(Advanced): Introduction to regex. Why and where to learn more

Checking for errors

1. **Reshaping and merging data using tidyr and dplyr**

The Tidy Data paradigm

Understanding the tidy data structure using R

Renaming and rearranging columns

Uses of the gather, spread, separate and unite functions

Putting data in order using arrange

Joining datasets by a common variables

(Advanced): Understanding different join types

(Advanced): Automating loading and tidying multiple files using plyr

1. **Mutating, filtering, arranging and summarising data using dplyr**

Mutating data: Adding columns without losing rows

Example of mutate

Example of transmute

(Advanced): Mutating many variables at once

Adding grouping variables (e.g. age categories)

Base R example using cut

Example using car’s ‘recode’ function

Mutating grouping variables using dplyr

Selecting, filtering and sorting data

Selecting columns

Filtering rows by a simple criterion

Filtering rows by multiple criteria

Arranging data by row values

Summarising data

Difference between mutating and summarising

The group\_by operation

Grouping by one variable (e.g. sex)

Grouping by two or more variables (e.g. sex and ethnicity)

Summarising a single variable

(Advanced) Summarising multiple variables at once using the funs() argument.

(Advanced) Introduction to data checking by visualising it in ggplot2

1. **Saving data using readr, readxl and haven**

Saving scripts

Saving tidy disaggregated data

Saving summary data

(Advanced) Saving as non-text files