



R Bootcamp

August 23-24, 2021

UBC
MFRE | Master of Food and
Resource Economics

Frequently used packages

- *We write packages in curly brackets – i.e. {dplyr} – and functions with () – i.e. mutate() – to differentiate them from each other*

Files / Reading data

- {[pacman](#)} – [p_load\(\)](#) to load and install packages
- {[here](#)} – [here\(\)](#) for file referencing
- {[readr](#)} – [read_csv\(\)](#) to read csv files
- {[readxl](#)} – [read_excel\(\)](#) to read xls/xlsx file

Frequently used packages

Data Wrangling

- `{dplyr}`
 - `mutate()` to add new variables, `select()` to pick variables, `filter()` to filter observations, `summarise()` to compute descriptive stats, `arrange()` to sort observations, `group_by()` to perform operations by group
- `{janitor}`
 - `clean_names()` to parse letter cases (capital or small letters) and separators (space, underscore, dash) of variable names to a consistent format

Frequently used packages

- **Statistics/Econometrics**

- {[stats](#)} – `lm()` for linear model, `t.test()` for student's t-test, `chisq.test()` for Pearson's chi-squared test, `cor()` for correlation
- {[lmtest](#)} – for diagnostic tests of regression models
- {[estimatr](#)} – `lm_robust()` for ordinary least squares with robust standard errors
- {[ivreg](#)} – `ivreg()` for two-stage least-squares regression

Frequently used packages

Formatting Results

- [{modelsummary}](#)
 - `modelsummary()` to format (rename, reorder, omit, etc) regression results
 - `datasummary()` to create data summaries, frequency tables, crosstabs, correlation tables, balance tables, etc.
- `{kable}` and `{kableExtra}` – [html](#) output, [pdf](#) output
 - `kbl()` to format tables into html or latex format
- [{stargazer}](#)
 - `stargazer()` to format tables into html or latex format

Frequently used packages

Data Visualization

- `{ggplot2}` – `ggplot()` for data visualization

Frequently used packages

Time Series Analysis

- [{xts}](#) – to convert data to time series object
- [{tsibble}](#) – to convert data to time series object that allows to use packages within the Tidyverse library (such as [{dplyr}](#))
- [{lubridate}](#) – working with dates
- [{forecast}](#) – for forecasting functions
- [{urca}](#) – for unit root and cointegration tests
- [{var}](#) – for VAR modelling

UBC MFRE

mfre.landfood.ubc.ca
