



STATLAB ACADEMY

Learn statistics by doing

R and Data Quick Start Guide

How to access R, import the Canadian macro dataset,
complete a first analysis, and solve common errors

Mohammad Safavi, Ph.D.

Student Guide | Version 1.0 | August 2026

statistics.safavim.com

1 Start in five minutes

The shortest reliable path from download to your first result.

Recommended for beginners: Posit Cloud

It runs RStudio in a web browser. You can begin without installing R on your computer. Use RStudio Desktop later if you work regularly, need offline access, or must keep data local.

- 1 Open statistics.safavim.com and choose Data Resources.
- 2 Find Canadian Macroeconomic Monthly Teaching Data and download the CSV, this guide, and the optional starter R script.
- 3 Open posit.cloud, sign in, and create a new RStudio project.
- 4 Upload the files. Keep the CSV beside your script, or keep the supplied data and scripts folders.
- 5 Run the starter script with the command below.

Run the complete starter analysis

```
source("scripts/START_HERE.R")
```

The script imports and validates the data, displays summaries, estimates a teaching-only regression, creates a PNG chart, and exports the latest 24 months. It uses base R only.

In this guide

Pages	What you will do
3-4	Download the correct files and choose Posit Cloud or RStudio Desktop.
5-7	Import, inspect, understand, and summarize the CSV.
8-10	Create charts, explore relationships, and export results.
11-12	Solve common errors and verify sources and limitations.

2 Get the right student files

Use the small Data Resources package, not the full textbook production archive.

On STATLAB, open Data Resources and locate the dataset card titled Canadian Macroeconomic Monthly Teaching Data. The site should offer three related downloads:

File	Purpose	Required?
canada_macro_monthly.csv	The actual teaching data used in R.	Yes
START_HERE.R	A tested first analysis that uses base R only.	Recommended
R and Data Quick Start Guide.pdf	This setup, analysis, and troubleshooting guide.	Recommended

Dataset at a glance

Coverage	Frequency	Rows	Variables	Missing
Jan 2017-Jun 2026	Monthly	114	5	None

If you download the complete student ZIP

Keep this folder structure

```

STATLAB_R_Data_Starter_Pack/
  README_FIRST.md
  data/
    canada_macro_monthly.csv
    DATA_DICTIONARY.md
    DATA_SOURCES.md
  scripts/
    START_HERE.R
  guide/
    STATLAB_R_and_Data_Quick_Start_Guide.pdf
  
```

Do not install this as an R package

This is a folder of course resources, not a CRAN package. Do not use `install.packages("STATLAB_R_Data_Starter_Pack")`. Extract the ZIP, open the folder as a project, and run the script.



3 Option 1: Posit Cloud

The easiest option when you want RStudio in a browser.

- 1 Go to <https://posit.cloud/> and sign in.
- 2 Create a new RStudio Project.
- 3 In the Files pane, select Upload and add the downloaded files or student ZIP.
- 4 If the ZIP remains compressed, extract it on your computer and upload the extracted folder contents.
- 5 Confirm that you can see the CSV. If using the complete pack, confirm that both data and scripts appear at the top level.
- 6 Open `scripts/START_HERE.R` and run it, or enter the source command in the Console.

```
source("scripts/START_HERE.R")
```

If you downloaded only the CSV

Upload the CSV to the top level of the project and enter:

```
macro <- read.csv("canada_macro_monthly.csv")  
macro$date <- as.Date(macro$date)  
View(macro)
```

Privacy reminder

Posit Cloud is convenient for public teaching data. For confidential, restricted, or personally identifiable data, follow your institution's rules and normally use an approved local or institutional environment.

Official help

Posit Cloud project documentation: docs.posit.co/cloud/guide/projects/

4 Option 2: RStudio Desktop

Best for regular, offline, private, or longer-running work.

- 1 Install the current stable version of R from cran.r-project.org. At publication, the current Windows release is R 4.6.1.
- 2 Install RStudio Desktop from posit.co/download/rstudio-desktop/.
- 3 Extract STATLAB_R_Data_Starter_Pack.zip to a normal folder. Do not work inside the compressed preview.
- 4 Open RStudio and choose File > New Project > Existing Directory.
- 5 Select the extracted STATLAB_R_Data_Starter_Pack folder.
- 6 Run the starter script.

```
source("scripts/START_HERE.R")
```

Which option should I use?

Need	Posit Cloud	RStudio Desktop
No installation	Best choice	No
Works offline	No	Best choice
Confidential data	Check policy first	Often preferable
Large or long analysis	Plan limits apply	Depends on computer
First class exercise	Best choice	Also works

No extra package is required

The supplied START_HERE.R script uses base R. Install ggplot2 only if you want to try the optional polished-chart example later in this guide.

5 Import and verify the CSV

Use reproducible code instead of relying only on menu clicks.

If you kept the supplied folder structure, run:

```
macro <- read.csv("data/canada_macro_monthly.csv")
macro$date <- as.Date(macro$date)

View(macro)
head(macro)
str(macro)
summary(macro)
```

If the CSV is beside your script, remove data/ from the path.

Run four checks

```
nrow(macro)           # expected: 114
ncol(macro)           # expected: 5
range(macro$date)     # 2017-01-01 to 2026-06-01
colSums(is.na(macro)) # expected: zero for every column
```

Preview of the first observations

date	CPI	inflation %	policy %	USD/CAD
2017-01-01	129.5	2.129	0.500	1.319276
2017-02-01	129.7	2.046	0.500	1.310726
2017-03-01	129.9	1.564	0.500	1.338643
2017-04-01	130.4	1.637	0.500	1.344021

Why convert the date?

`read.csv()` initially reads the ISO date as text. `as.Date()` converts it to an R Date, allowing correct sorting, filtering, plotting, and time calculations.

6 Know your variables

Read the units before calculating, graphing, or interpreting.

Variable	Unit	Plain-language meaning
date	Month	First calendar day identifying each observation month.
cpi_all_items_2002_100	Index, 2002 = 100	Canada all-items Consumer Price Index level.
inflation_yoy_percent	Percent	Twelve-month percentage change in the CPI.
policy_rate_percent	Percent	Last available target for the overnight rate in the month.
usd_cad_monthly_average	CAD per USD	Average Canadian-dollar cost of one US dollar.

Three distinctions that prevent common mistakes

- CPI level is not inflation. The CPI is an index; inflation is its rate of change.
- USD/CAD direction matters. A higher value means one US dollar costs more Canadian dollars.
- Frequency matters. CPI is monthly, while the source policy and exchange-rate series are daily and are transformed to monthly values.

Useful selection examples

```
# Keep observations from 2024 onward
recent <- subset(macro, date >= as.Date("2024-01-01"))

# Select three variables
analysis_data <- macro[c("date", "inflation_yoy_percent",
                        "policy_rate_percent")]

# Latest observation
tail(macro, 1)
```

Frozen teaching data

The file is intentionally fixed at Version 1.0 so every student can reproduce the same output. It does not update automatically when official agencies revise or extend a series.

7 Complete your first analysis

Summarize inflation and then communicate the pattern visually.

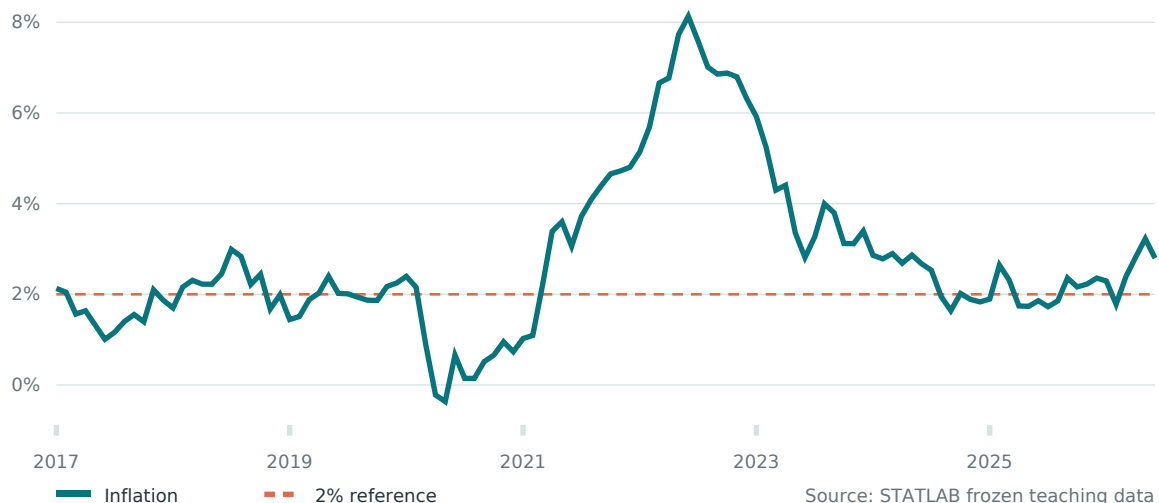
```
# Descriptive statistics
mean(macro$inflation_yoy_percent)
median(macro$inflation_yoy_percent)
sd(macro$inflation_yoy_percent)
range(macro$inflation_yoy_percent)

# Latest month
tail(macro[c("date", "inflation_yoy_percent")], 1)
```

Check	Expected Version 1.0 result
Mean inflation	2.774%
Minimum inflation	-0.366%
Maximum inflation	8.133%
Latest observation	June 2026: 2.798%

Canadian year-over-year inflation

January 2017 to June 2026



The line shows the historical path in the frozen sample. The dashed 2% line is a visual reference, not a forecast or trading signal.

8 Explore relationships responsibly

Correlation and regression are useful summaries, but they do not prove cause and effect.

Start with a correlation matrix:

```
variables <- macro[c("inflation_yoy_percent",  
                    "policy_rate_percent",  
                    "usd_cad_monthly_average")]  
  
round(cor(variables), 3)
```

Then estimate a simple teaching model:

```
model <- lm(  
  inflation_yoy_percent ~ policy_rate_percent +  
    usd_cad_monthly_average,  
  data = macro  
)  
  
summary(model)  
confint(model)
```

How to read the output

- The Estimate is the fitted sample association, holding the other included variable constant.
- The Std. Error, t value, and p-value summarize uncertainty under the model assumptions.
- The R-squared reports the share of sample variation accounted for by the fitted linear model.
- Time-series observations can be autocorrelated. A basic contemporaneous regression may therefore require better dynamics and standard errors for serious inference.

Do not make a causal claim

Policy rates respond to inflation and other economic conditions. Exchange rates and inflation can also move together for many reasons. This model is for learning R syntax and output interpretation, not identifying policy effects.

9 Create a clear time-series graph

Begin with base R; use ggplot2 only when you want the optional extension.

Base R version - no installation required

```
plot(
  macro$date, macro$inflation_yoy_percent,
  type = "l", lwd = 2.5, col = "#08747D",
  xlab = "Month",
  ylab = "Year-over-year inflation (%)",
  main = "Canadian inflation"
)
abline(h = 2, lty = 2, col = "#D96C4B")
```

Optional ggplot2 version

Install once per computer or cloud project, then load it in each new R session:

```
install.packages("ggplot2") # run once only
library(ggplot2)

ggplot(macro, aes(date, inflation_yoy_percent)) +
  geom_line(linewidth = 0.9, colour = "#08747D") +
  geom_hline(yintercept = 2, linetype = 2,
            colour = "#D96C4B") +
  labs(
    title = "Canadian inflation",
    x = NULL, y = "Year-over-year inflation (%)",
    caption = "STATLAB frozen teaching data"
  ) +
  theme_minimal(base_size = 12)
```

Graph honestly

Label the measure and unit, keep dates in chronological order, identify the source, and avoid a truncated scale when it would exaggerate small changes. Explain any reference line.

10 Save your work

Export a clean table, a figure, and a reusable script.

Export the latest 24 months

```
recent <- tail(macro, 24)
write.csv(recent, "canada_macro_last_24_months.csv",
          row.names = FALSE)
```

Save a base R plot as a PNG

```
png("canadian_inflation_plot.png",
    width = 1400, height = 800, res = 160)
plot(macro$date, macro$inflation_yoy_percent,
     type = "l", lwd = 2.5, col = "#08747D",
     xlab = "Month",
     ylab = "Year-over-year inflation (%)")
abline(h = 2, lty = 2, col = "#D96C4B")
dev.off()
```

Run the supplied starter script

```
source("scripts/START_HERE.R")
```

It creates both files automatically in the active project folder. Find them in RStudio's Files pane and select More > Export or download them through your browser.

A reproducible workflow

Step	Good practice
1. Question	Write what you want to learn before choosing a method.
2. Data	Keep the untouched CSV and read the dictionary and source notes.
3. Code	Save commands in an R script rather than only in Console history.
4. Output	Export tables and figures with clear filenames.
5. Interpretation	State the finding, uncertainty, and limitation together.

Keep inputs separate from outputs

Do not overwrite the original CSV. Save cleaned or selected data under a new name so your analysis can always be rebuilt from the frozen input.

11 Solve common problems

Use these checks before changing code or reinstalling software.

Message or symptom	What it means	Fix
cannot open file	R is looking in the wrong folder or the filename differs.	Run <code>getwd()</code> and <code>list.files(recursive = TRUE)</code> . Use the path shown.
object 'macro' not found	The import command did not run successfully.	Run <code>read.csv()</code> first and resolve any file error.
could not find function <code>read_macro</code>	That helper belongs to the full book source, not this starter pack.	Use <code>read.csv()</code> directly or run <code>START_HERE.R</code> .
there is no package called <code>ggplot2</code>	The optional package is not installed in this project.	Run <code>install.packages("ggplot2")</code> once.
<code>scripts/START_HERE.R</code> typed as code	A filename alone is not an R command.	Use <code>source("scripts/START_HERE.R")</code> .
Dates sort incorrectly	The date column is still text.	Run <code>macro\$date <- as.Date(macro\$date)</code> .
ZIP opens but code fails	Files are still inside a compressed preview.	Extract the ZIP and open the extracted folder as the project.

Diagnostic commands

```
getwd()
list.files()
list.files(recursive = TRUE)
file.exists("data/canada_macro_monthly.csv")
sessionInfo()
```

Avoid `setwd()` in shared course code

A path such as `C:/Users/Name/Desktop` works only on one computer. Open the extracted folder as an RStudio project and use project-relative paths such as `data/canada_macro_monthly.csv`.

12 Sources, limits, and next steps

Know where the numbers came from and what they can support.

Official sources

- Statistics Canada: Table 18-10-0004-01, Canada all-items CPI, vector v41690973. [Official bulk CSV](#) and [Open Licence](#).
- Bank of Canada: Valet series V39079 for the target overnight rate and FXUSDCAD for the daily USD/CAD average. [Valet documentation](#) and [terms of use](#).

How the monthly file was constructed

- CPI remains at monthly frequency.
- Year-over-year inflation equals $100 \times (\text{CPI in month } t / \text{CPI in month } t-12 - 1)$.
- The policy rate is the last available daily target in each month.
- USD/CAD is the arithmetic mean of available daily averages in each month.
- The file was frozen on August 11, 2026. Later official revisions may differ.

Responsible interpretation

This dataset and guide are for education. They do not provide investment, trading, legal, accounting, or policy advice. Historical associations are not forecasts, and a simple regression is not evidence of causality.

Continue learning

Return to the Introduction to R module on [STATLAB Academy](#). Work through each example in the sequence: question, data, R code, output, interpretation, and limitation.



STATLAB Academy

Statistics made visual, practical, and reproducible.

<https://statistics.safavim.com>