

MIEF Skills Workshop: Modeling in R

Session 1

January 30, 2026

Course Purpose

What this course will do:

- Introduce you to iterative regression methods in R
- Provide you with a template to conduct statistical modeling, analysis, and visualization in R

What this course WON'T do:

- Teach you econometrics

Today

This course will involve less theory than the previous course. You should refer to the introductory R course if you're fuzzy on basic R details.

This course will be far more **practical**. I will be building an R script alongside you, as you build your own R script to "learn by doing."

Today, we will be:

- Building a simple statistical model to answer a very simple research question: one dependent variable, one independent variable
- Fleshing the statistical process out more with control variables and fixed effects
- Iterating our statistical process to encompass multiple dependent variables
- Exploring a simple method to output regression tables
- Iterating that method to work for multiple dependent variables

Quick Refresh: Tidyverse

Base R Layout

Tidyverse Layout

```
names(iris)
```

```
## [1] "Sepal.Length" "Sepal.Width"    "Petal.Length"  "Petal.Width"   "Species"
```

```
str_replace(str_to_lower(names(iris)), "\\.", "_")
```

```
## [1] "sepal_length" "sepal_width"   "petal_length" "petal_width"  "species"
```

Quick Refresh: Tidyverse

Base R Layout

Tidyverse Layout

Tidyverse functions introduce a 'cleaner' method to write code out, using what is called the 'pipe operator': `%>%`. It's almost like writing a recipe, step by step.

```
names(iris)
```

```
## [1] "Sepal.Length" "Sepal.Width"    "Petal.Length" "Petal.Width"   "Species"
```

```
iris %>%  
  names() %>%  
  str_to_lower() %>%  
  str_replace("\\"., "_")
```

```
## [1] "sepal_length" "sepal_width"   "petal_length" "petal_width"  "species"
```

(Note — you can also write the pipe operator as `|>`)

Quick Refresh: Using `map()`

The `map()` function from the `purrr` package allows us to **apply the same function(s) to a set of objects**.

```
library(purrr)
purrr::map(
  .x = list(
    c(-10, 0, 10), c(1, 2, 3, 4, 5)
  ),
  .f = function(x)
    (x - min(x, na.rm = TRUE)) /
    (max(x, na.rm = TRUE) - min(x, na.rm = TRUE))
)
```

```
## [[1]]
## [1] 0.0 0.5 1.0
##
## [[2]]
## [1] 0.00 0.25 0.50 0.75 1.00
```

Quick Refresh: Using `map()`

`map()` has a shorthand to simplify its use: instead of writing the function out, you can replace `function() {}` with `\(argument)` and the function's argument with `argument` ("argument" can be replaced with whatever you want).

```
c(1, 3) %>%  
  map(  
    function(x) {  
      x + 10  
    }  
  ) %>%  
  as.numeric()
```

```
## [1] 11 13
```

```
c(1, 3) %>%  
  map(  
    \vlu vlu + 10  
  ) %>%  
  as.numeric()  
  
## [1] 11 13
```

Setup

At this point, you should know how to set up a data project folder. Do this now. This folder should include:

- A "data" folder
- A "code" folder
- An "output" folder
- A **.rproj** file

Download the course data and documentation from Canvas. Once this is done, **open a new R script**. I will go through a chunk of code, and then you will be expected to replicate it.

The World Values Survey

"The survey, which started in 1981, seeks to use the most rigorous, high-quality research designs in each country. The WVS consists of nationally representative surveys conducted in almost 100 countries which contain almost 90 percent of the world's population, using a common questionnaire. [...] WVS seeks to help scientists and policy makers understand changes in the beliefs, values and motivations of people throughout the world."

- Social values, attitudes & stereotypes
- Societal well-being
- Social capital, trust and organizational membership
- Economic values
- Corruption
- Migration
- Post-materialist index
- Science & technology
- Religious values
- Security
- Ethical values & norms
- Political interest and political participation
- Political culture and political regimes
- Demography

Today's Practical Component

- You will be following my live coding and then replicating it on your script.
- After each coding chunk, you will get 5-10 minutes to apply it to your script:
 - If you are not feeling confident in the live setting, you can exactly replicate what I am doing
 - If you would like to get a head start on your course assignment, you can apply the process I am following to **a different set of variables within the dataset**.

Today's Practical Component

I am answering the research question: what relationship is there between a respondent's education level and how they raise their children? To do this, I am using Q275 (education level) as the independent variable, and Q7-17 (child values) as the dependent variables from the World Values Survey.

For the purpose of today's session, I would recommend sticking to education level as the independent variable. Here are some other dependent variables you could consider using:

- Life-Importance Subjects (Q1-6)
- Neighbor Priorities (Q18-26)
- Trust in Other People (Q58-63)
- Confidence in Institutions (Q64-89, NOT with 82_suffix options)
- Organization Membership (Q94-104, NOT with _r suffix)
- Corruption Perception (Q113-117)
- Immigration Perception (Q122-129)
- Information Source (Q201-208)
- Elections Perception (Q224-233)

Today's Practical Component

1. Setup
2. Import Data
3. Simple Regression – One Dependent Variable
4. Iterated Regressions – Multiple Dependent Variables
5. Simple Visualization (Stargazer and LateX)
6. Iterated Visualizations (Stargazer and LateX)
7. Outputting Iterated Regressions as a Dataset