Multilevel Visualization

Andrew Grogan-Kaylor

2024-04-19

Table of contents

List of Figures

List of Tables

1 Multilevel Visualization

> "Persist and verify… The power that we abdicate to others out of our insecurity - to others who insult us with their faux-intuition or their authoritarian smugness - that comes back to hurt us so deeply… But the power we wrest from our own certitude - that saves us." (Cash 2017)

> "Mathematical Science shows us what is. It is the language of unseen relations between things. But to use & apply that language we must be able fully to appreciate, to feel, to seize, the unseen, the unconscious. Imagination too shows us what is, the is that is beyond the senses." (Lovelace 1992)

1.1 Introduction

Below, I describe the use of Stata (StataCorp 2021), R[1] (R Core Team 2023), and Julia (Bezanson et al. 2017) to visualize multilevel models.

1.2 The Data

The examples use the `simulated_multilevel_data.dta` file from *Multilevel Thinking*. Here is a direct link to download the data.

Table 1.1: Sample of Simulated Multilevel Data

Table 1.1: Table continues below

country	HDI	family	id	identity	intervention	physical_punishment
1	69	1	1.1	2	1	3
1	69	2	1.2	2	2	2
1	69	3	1.3	1	2	3
1	69	4	1.4	2	1	0
1	69	5	1.5	2	1	4
1	69	6	1.6	1	2	5

[1] In R, I use the `ggplot2` (Wickham 2016) library.

Table 1.2: Sample of Simulated Multilevel Data

warmth	outcome
3	58.47
1	51.1
2	53.92
5	61.17
4	56.05
3	50.81

2 Graphs

2.1 Scatterplots

A scatterplot is one of the most basic of all data visualizations. At the same time, a scatterplot can be tremendously informative because it provides: the location of every data point (data points may be overprinted); a sense of the distribution of both the x and y variables; and a sense of the overall trend in the relationship between the two variables, if there is one.

2.1.1 Stata

2.1.1.1 Get The Data

```
use simulated_multilevel_data.dta
```

2.1.1.2 Scatterplot

```
twoway scatter outcome warmth, ///
  xtitle("warmth") ytitle("outcome") ///
  title("Outcome by Parental Warmth")

quietly graph export scatter.png, replace
```

2.1.2 R

2.1.2.1 Get The Data

```
library(haven)

df <- read_dta("simulated_multilevel_data.dta")
```

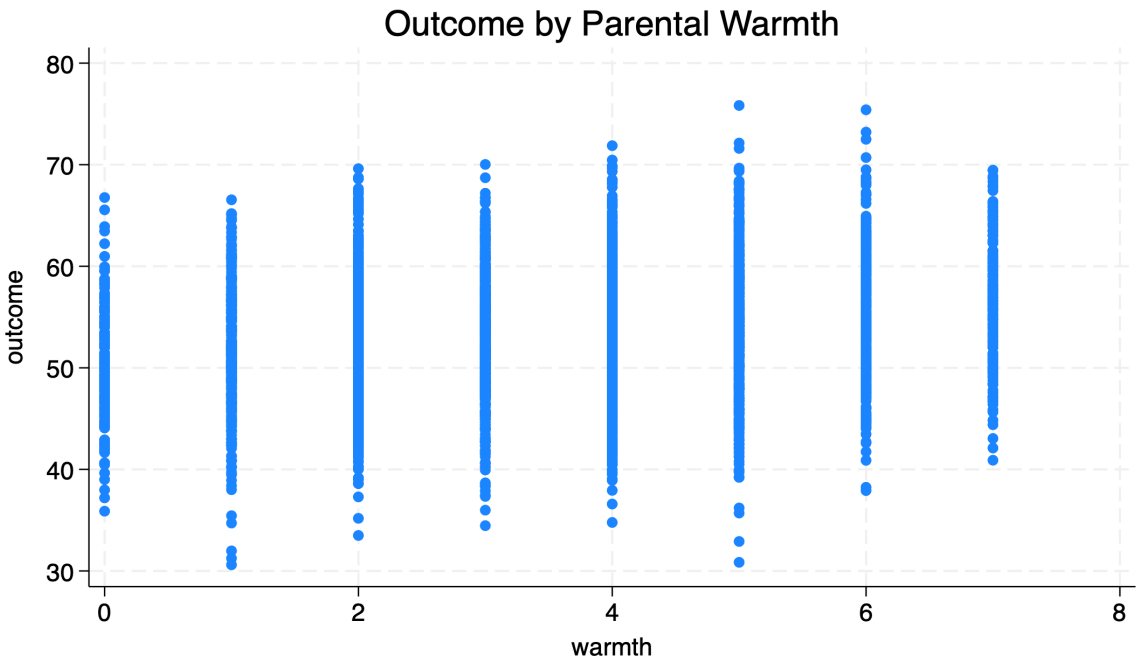

Figure 2.1: Outcome by Parental Warmth (Stata)

2.1.2.2 Scatterplot

```
library(ggplot2)

ggplot(df,
       aes(x = warmth,
           y = outcome)) +
  geom_point() +
  labs(title = "Outcome by Parental Warmth")
```

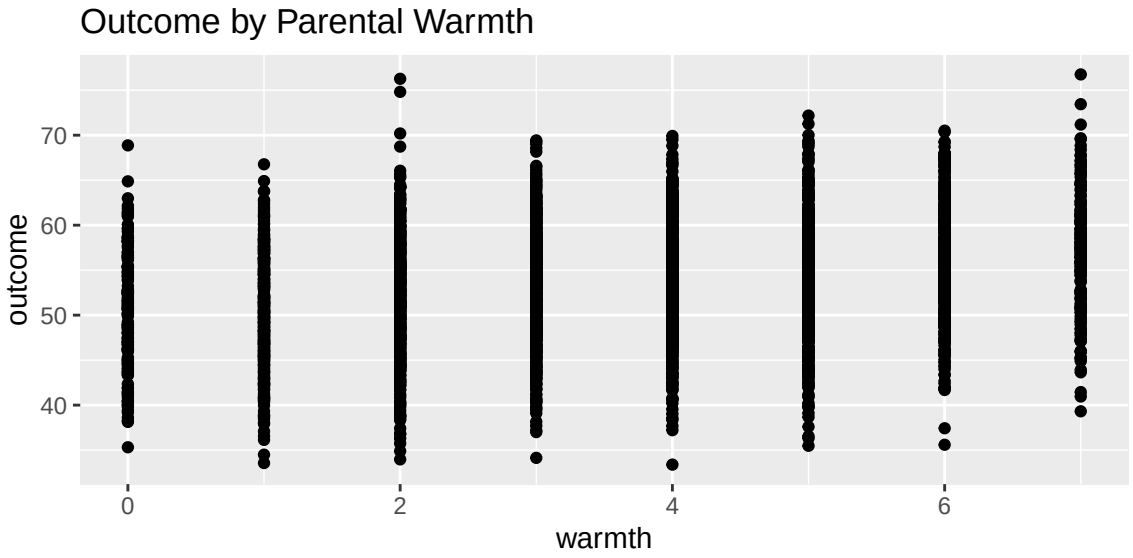

Figure 2.2: Outcome by Parental Warmth (R)

2.1.3 Julia

2.1.3.1 Get The Data

```
using Tables, MixedModels, StatFiles, DataFrames, CategoricalArrays, DataFramesMeta

df = DataFrame(load("simulated_multilevel_data.dta"))
```

2.1.3.2 Scatterplot

```
using StatsPlots

@df df scatter(:warmth, :outcome,
               title = "Outcome by Parental Warmth",
               ylabel = "outcome",
               xlabel = "parental warmth")
```

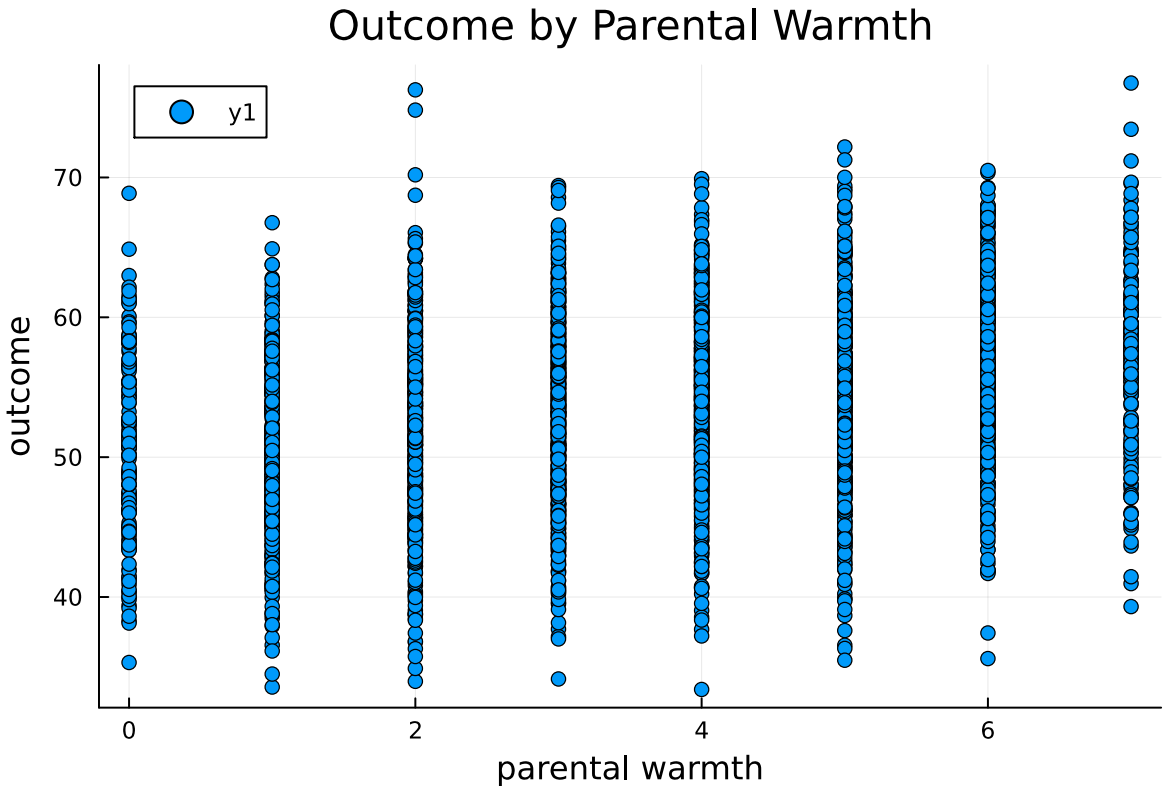

Figure 2.3: Outcome by Parental Warmth (Julia)

2.2 Line Graph (Linear Trend)

A line graph of the data focuses in on the linear trend in the data.

2.2.1 Stata

2.2.1.1 Get The Data

```
use simulated_multilevel_data.dta
```

2.2.1.2 Line Graph

```
twoway lfit outcome warmth, ///
  xtitle("warmth") ytitle("outcome") ///
  title("Outcome by Parental Warmth")

quietly graph export lfit.png, replace
```

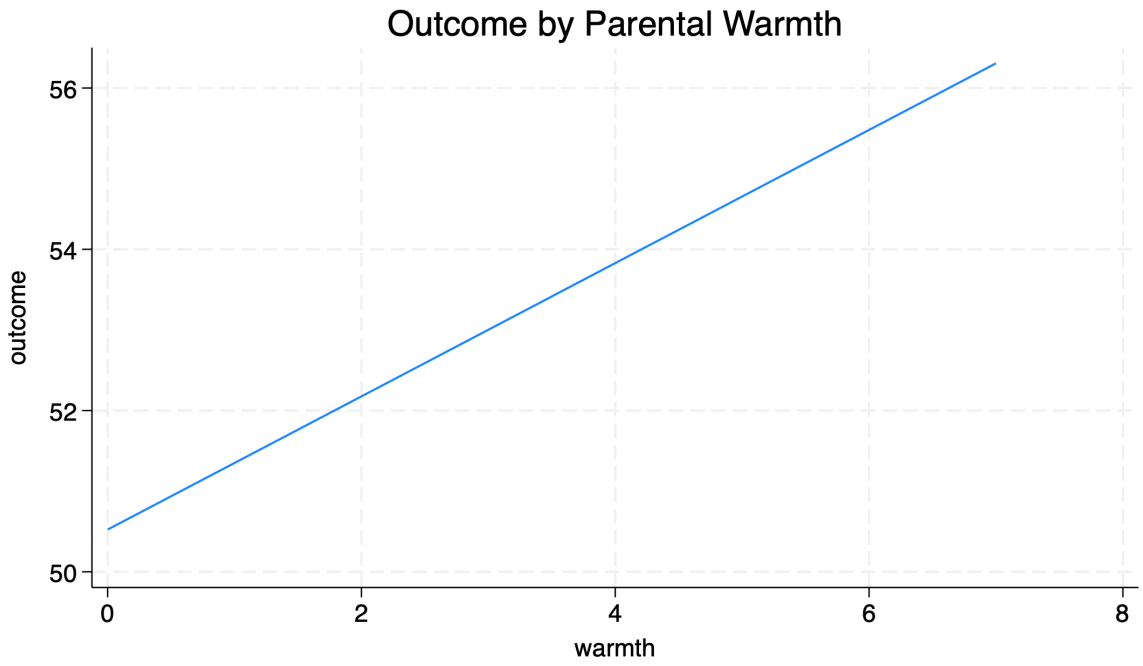

Figure 2.4: Outcome by Parental Warmth (Stata)

2.2.2 R

2.2.2.1 Get The Data

```
library(haven)

df <- read_dta("simulated_multilevel_data.dta")
```

2.2.2.2 Line Graph

```
library(ggplot2)

ggplot(df,
       aes(y = outcome,
           x = warmth)) +
  geom_smooth(method = "lm") +
labs(title = "Outcome by Parental Warmth")
```

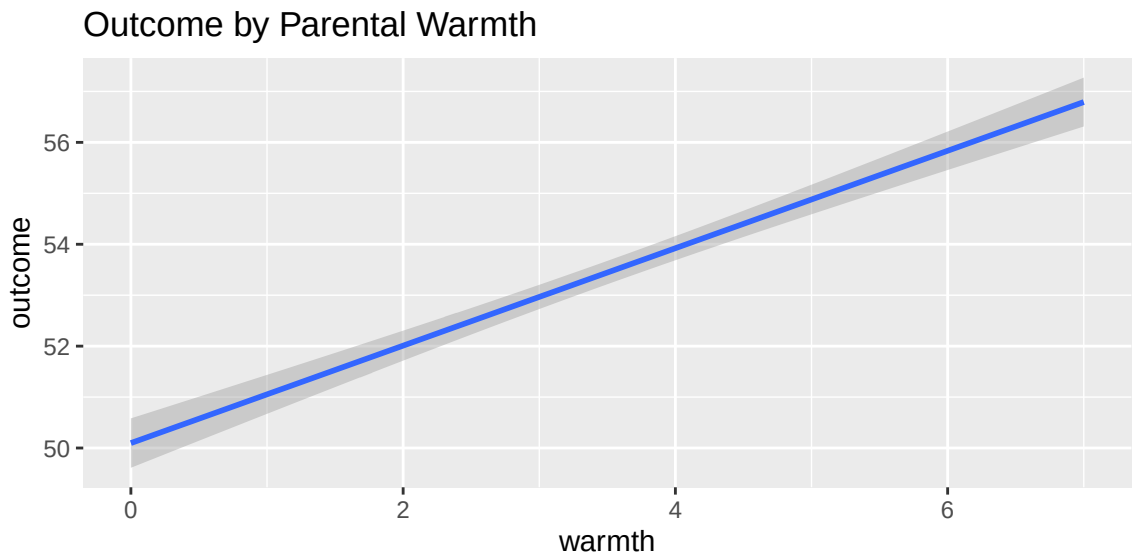

Figure 2.5: Outcome by Parental Warmth (R)

2.2.3 Julia

2.2.3.1 Get The Data

```
using Tables, MixedModels, StatFiles, DataFrames, CategoricalArrays, DataFramesMeta

dfL = DataFrame(load("simulated_multilevel_data.dta"))
```

2.2.3.2 Line Graph

To make our plot with a smoother in Julia, we set the `markercolor` and `markerstrokecolor` to be *white*, and the `smooth` option to `:true`.

```
using StatsPlots

@df dfL scatter(:warmth, :outcome,
                title = "Outcome by Parental Warmth",
                ylabel = "outcome",
                xlabel = "warmth",
                markercolor = "white",
                markerstrokecolor = "white",
                smooth=:true)
```

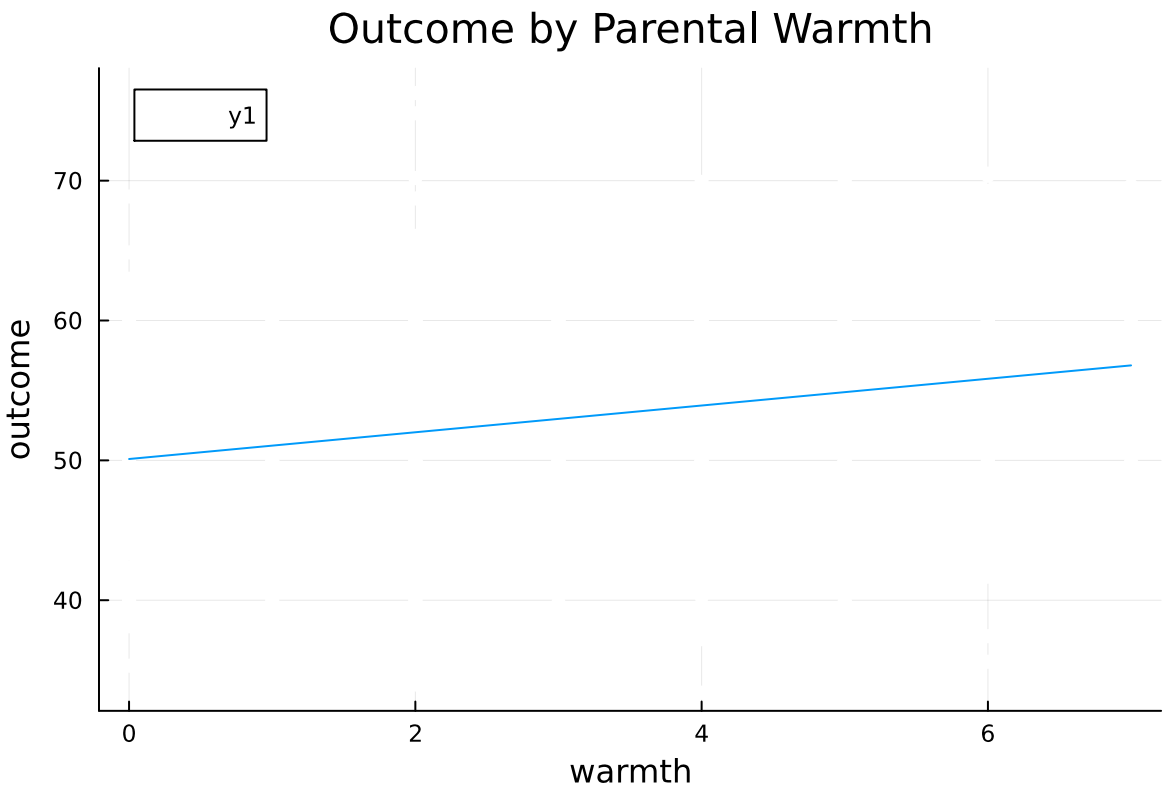

Figure 2.6: Outcome by Parental Warmth (Julia)

2.3 Spaghetti Plots

A *spaghetti plot* might be considered the most *multilevel* of the visualizations here considered. A spaghetti plot shows the group specific slopes and intercepts for all of the groups in the data.

2.3.1 Stata

In Stata, spaghetti plots are most easily generated using the user written `spagplot` command. Type `findit spagplot` to install this command.

2.3.1.1 Get The Data

```
use simulated_multilevel_data.dta
```

2.3.1.2 Spaghetti Plot

```
spagplot outcome warmth, ///
  id(country) ///
  xtitle("parental warmth") ytitle("outcome") ///
  title("Outcome by Parental Warmth")

quietly graph export spagplot.png, replace
```

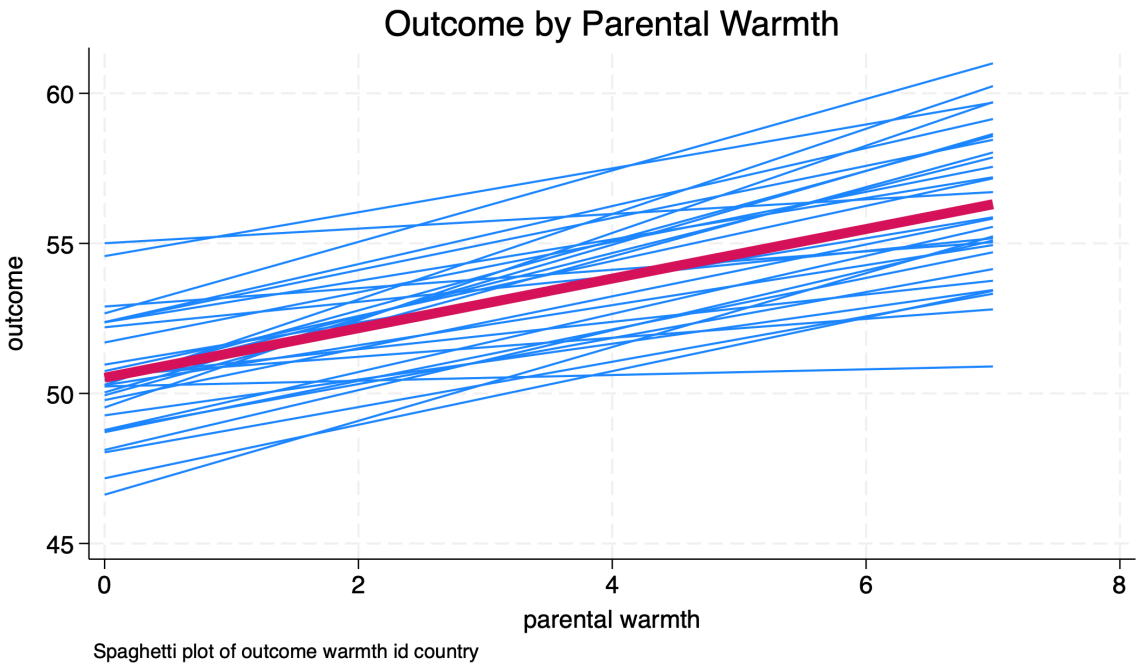

Figure 2.7: Outcome by Parental Warmth (Stata)

2.3.2 R

2.3.2.1 Get The Data

```
library(haven)

df <- read_dta("simulated_multilevel_data.dta")
```

2.3.2.2 Spaghetti Plot

```
library(ggplot2)

df$country <- factor(df$country)

ggplot(df,
       aes(y = outcome,
           x = warmth)) +
  geom_smooth(aes(color = country,
                  group = country),
              method = "lm",
              se = FALSE) +
    geom_smooth(method = "lm", linewidth = 3) +
labs(title = "Outcome by Parental Warmth")
```

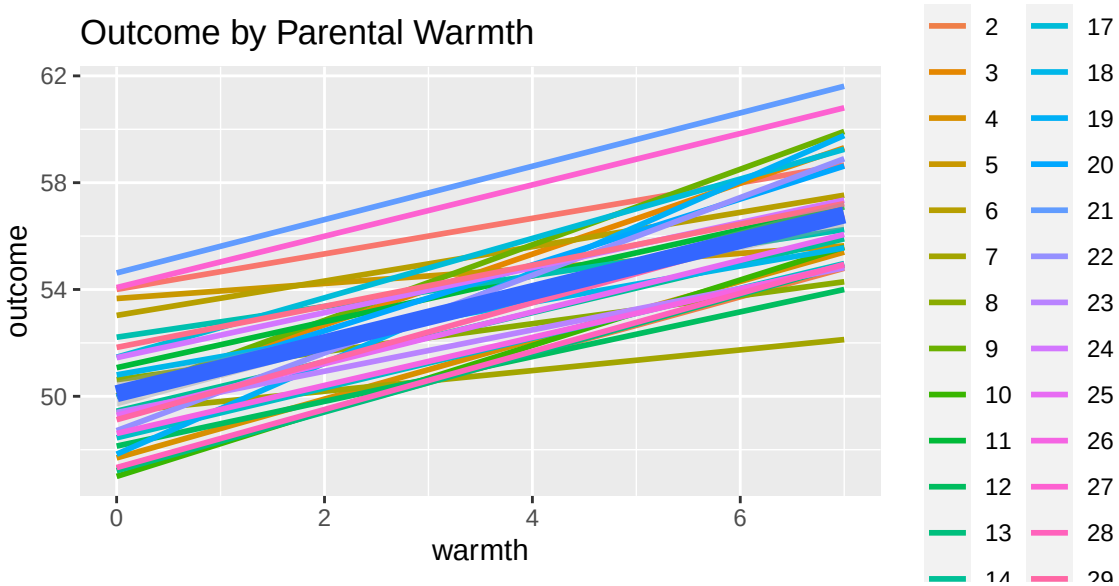

Figure 2.8: Outcome by Parental Warmth (R)

2.3.3 Julia

2.3.3.1 Get The Data

```
using Tables, MixedModels, StatFiles, DataFrames, CategoricalArrays, DataFramesMeta

dfL = DataFrame(load("simulated_multilevel_data.dta"))
```

2.3.3.2 Spaghetti Plot

```
using StatsPlots

@df dfL scatter(:warmth, :outcome,
                title = "Outcome by Parental Warmth",
                ylabel = "outcome",
                xlabel = "warmth",
                markercolor = "white",
                markerstrokecolor = "white",
                group = :country,
```

```
                legend = false,
                smooth=:true)
```

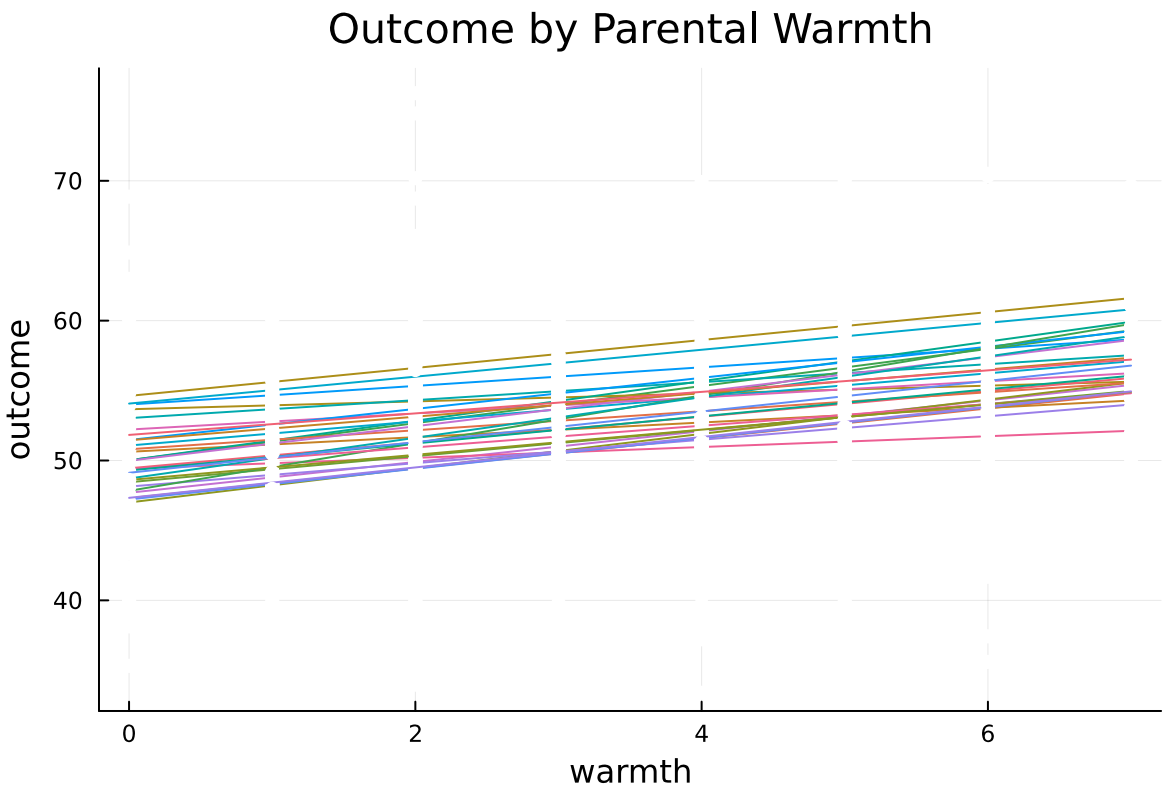

Figure 2.9: Outcome by Parental Warmth (Julia)

References

Bezanson, Jeff, Alan Edelman, Stefan Karpinski, and Viral B. Shah. 2017. "Julia: A Fresh Approach to Numerical Computing." *SIAM Review* 59 (1): 65–98. https://doi.org/10.1137/141000671.

Cash, Roseanne. 2017. "Roseanne Cash Reads 'Power' by Adrienne Rich." In *The Universe in Verse*.

Lovelace, Ada King. 1992. *Ada: The Enchantress of Numbers: A Selection from the Letters of Lord Byron's Daughter and Her Description of the First Computer*. Edited by Betty A. Toole. Strawberry Press.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

StataCorp. 2021. *Stata 17 Graphics Reference Manual*. Stata Press.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.