

# Final Project

Yigit Tahmisoglu

2023-04-01

Github repository: [https://github.com/tahmisoglu-yigit/R\\_Project.git](https://github.com/tahmisoglu-yigit/R_Project.git)

Requirements: 1. 1-page executive summary (can include 1 graph or image) 4. Data exploration, questions you tried to answer, interesting things. This should be similar to the data exploration in R for Data Science (Chapter 7) and most other parts of the book. No hard results needed. (3-5 pages) 5. Code: Include one piece of code that you describe in some detail what it does. Pick your favorite graph or table, and explain the choices you made and what each step does. A pipeline ( $|>$ ) of 4 to 7 steps is sufficient. The goal is to have you write and explain code.

## Summary of the raw data set

The data set is the one I plan to use for my masters thesis. It is a cross-sectional survey data from the Demographic and Health Surveys for Turkey from the year 2018, which is a representative household survey providing detailed information on birth recodes and individual records for women of each household for developing countries. The potential questions I am interested in exploring are: - Effect of childbirth on female labor force participation

**Explain how you cleaned the data, or if it was clean, how you ensured that it was. Clearly mention any R scripts that did the checking. (It should be automated, that means no “I printed it a few times to the screen and stared at it.”) (1-2 pages, less if data was clean) :**

Install libraries:

Variable names are not descriptive. But taking the DHS Manual as reference, we can rename the variable of interests to make further analyses easier.

Also rename unique values of categorical variables: