Life expectancy*

Yufei Liu

October 31, 2023

1 Introduction

2 Data

We used `R` to do the analysis in this paper (R Core Team 2022). We used packages `tidyverse`(Wickham et al. 2019), `janitor`(Firke 2021), and `here`(Müller 2020) to clean and load the data as well as create figures, and `knitr`(Xie 2014) and `modelsummary`(**citeModelsummary?**) to generate tables. The color style of the figures was created referring to a R colors cheet-sheet (**citeRcolor?**).

2.1 Data description

The datasets used in this paper was obtained from World Health Organization (WHO) Global Health Observatory data repository, and is publicly available from WHO website (**citeWHO?**). We utilized 3 datasets: life expectancy and Health life expectancy by WHO region(**citeLife?**), SDG Target 3.5 Substance abuse by WHO region (**citeAlcohol?**), and SDG Target 3.a Tobacco control by WHO region (**citeTobacco?**). The life expectancy dataset contains life expectancy and helthy life expectancy at birth and at age 60 for different sexes and WHO regions in year 2000, 2010, 2015, and 2019. The substance abuse data contains

*Code and data are available at: https://github.com/Florence-Liu/life_expectancy

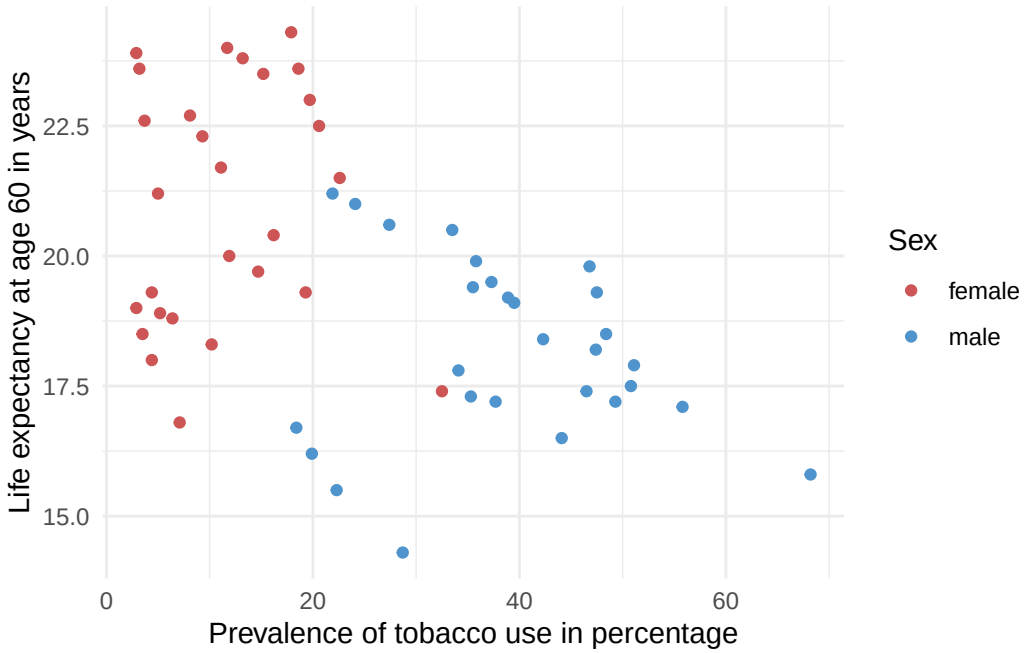

Figure 1: The effect of prevalence of tobacco on life expectancy at age 60 for different sex

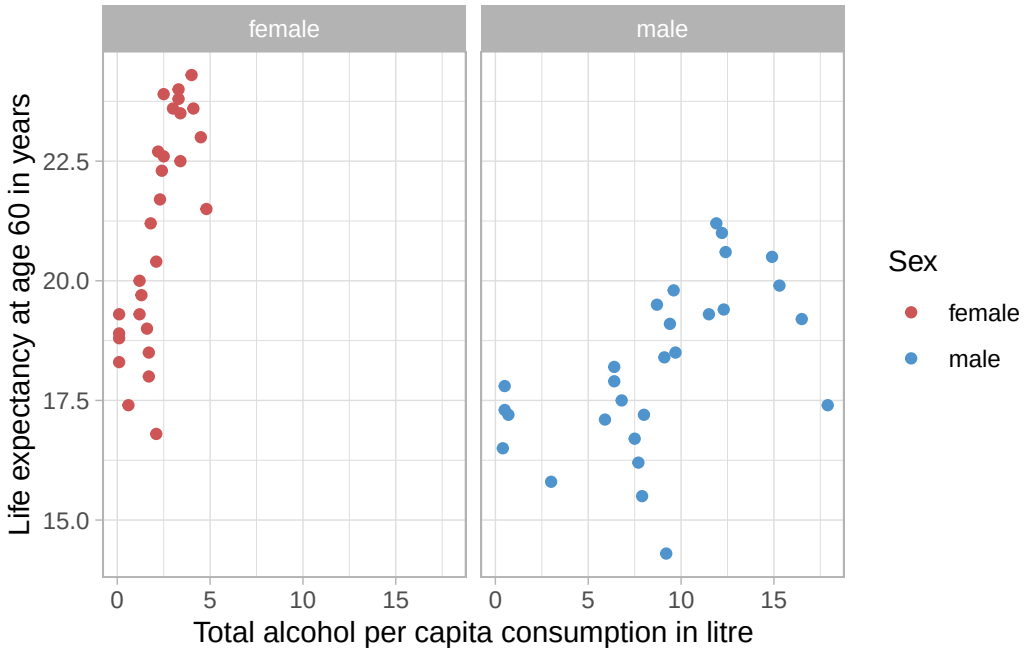

Figure 2: The effect of total alcohol per capita consumption on life expectanct at age 60 for different sex

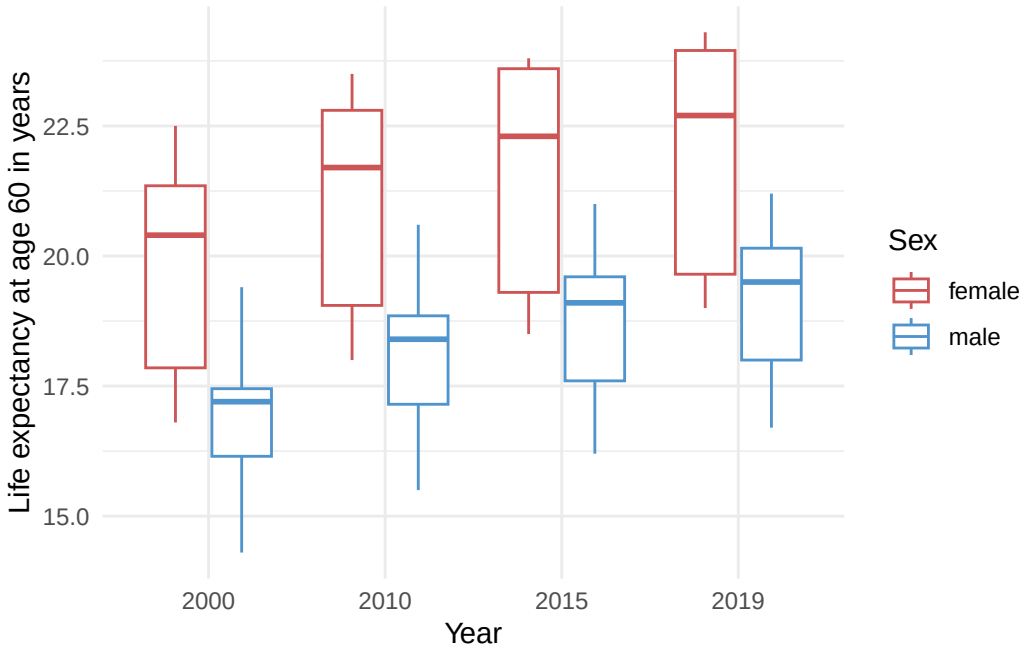

Figure 3: Life expectancy at age 60 across years for different sex

Table 1: Summary of average life expectancy, prevalence of tobacco, and alcohol consumption across years

Year	Mean life expectancy (year)	Mean prevalence of tobacco (%)	Mean alcohol consumption (litre)
2000	18.3	31.2	5.2
2010	19.5	25.2	5.6
2015	20.1	22.9	5.6
2019	20.5	21.4	5.3

Table 2: Summary of average life expectancy, prevalence of tobacco, and alcohol consumption for different WHO regions

Region	Mean life expectancy (year)	Mean prevalence of tobacco (%)	Mean alcohol consumption (litre)
Africa	16.9	13.4	4.9
Americas	22.0	21.2	7.8
Eastern Mediterranean	18.0	22.2	0.3
Europe	21.2	29.2	10.2
Global	20.2	26.6	5.5
South-East Asia	18.2	37.6	3.2
Western Pacific	20.8	26.0	5.9

3 Model

4 Result

Table 3: Summary of two linear regression models

	Tobacco	Alcohol*Sex
(Intercept)	21.401	18.165
	(0.528)	(0.557)
prevalence_of_tobacco	−0.072	
	(0.017)	
alcohol_consumption		1.302
		(0.217)
sexmale		−1.686
		(0.823)
alcohol_consumption × sexmale		−1.106
		(0.226)
Num.Obs.	56	56
R2	0.238	0.646
R2 Adj.	0.224	0.625
AIC	250.9	212.0
BIC	256.9	222.1
Log.Lik.	−122.425	−100.976
F	16.852	31.592
RMSE	2.15	1.47

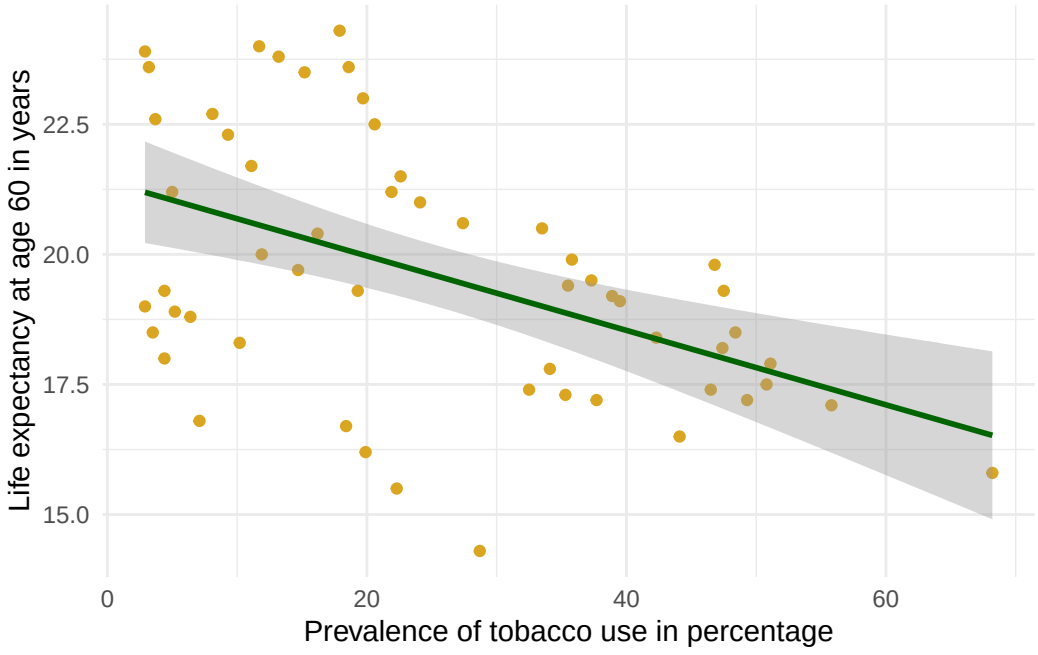

Figure 4: Prevalence of tobacco use versus life expectancy at age 60 with fitted regression line

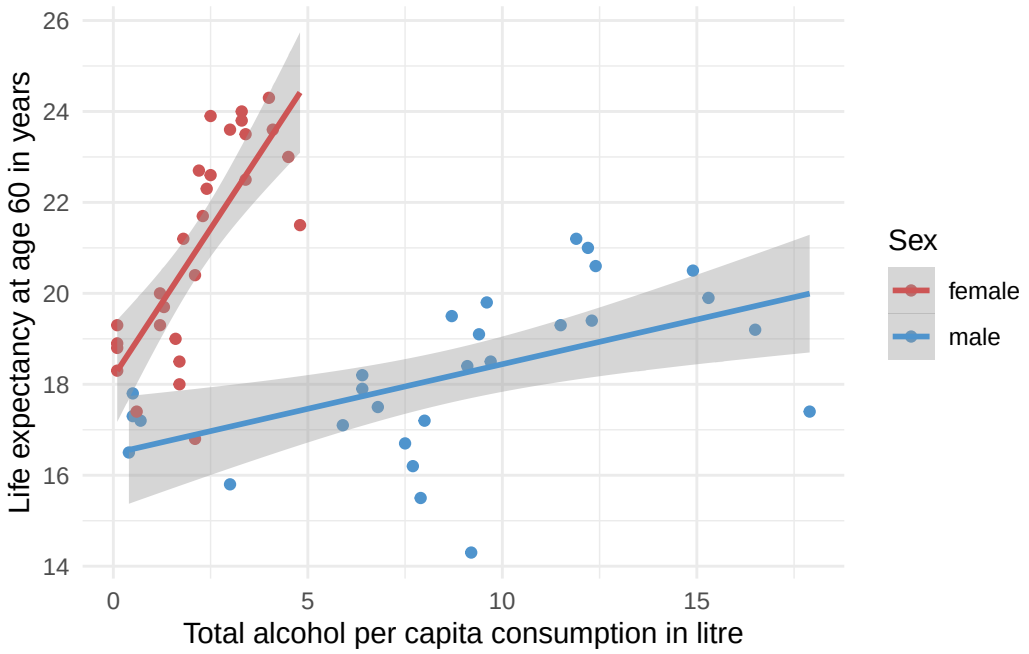

Figure 5: Alcohol consumption per capita versus life expectancy with fitted regression line for different sex group

Reference

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.