Factors influencing Total Income in the United States*

Research and Statistical Analysis on Factors influencing Total Income in the United States of America

Wen Han Zhao

December 1, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

Table of contents

---

*Code and data are available at: https://github.com/younazhao/Income-Influencing-Factors.git.

1

1 Introduction

[] for citation appendix WEIGHT The weight (X42001) is a partially design-based weight constructed at the Federal Reserve using original selection probabilities and frame information along with aggregate control totals estimated from the Current Population Survey. This weight is a relatively minor revision of the consistent weight series (X42000) maintained for the SCFs beginning with 1989 (For a detailed discussion of these weights, see "Consistent Weight Design for the 1989, 1992, and 1995 SCFs and the Distribution of Wealth," by Arthur B. Kennickell and R. Louise Woodburn, Review of Income and Wealth, Series 45, Number 2, June 1999, pp. 193-215 or the longer version given on the SCF web site at www.federalreserve.gov/econresdata/scf/scf_workingpapers.htm). The nature of the revisions to the consistent weights is described in "Revisions to the SCF Weighting Methodology: Accounting for Race/Ethnicity and Homeownership" (Arthur B. Kennickell, January 1999, www.federalreserve.gov/econresdata/ scf/scf_workingpapers.htm). A version of the revised weight has been computed for all the surveys beginning with 1989, and this variable has been added to the public versions of the SCF data sets. Users should be aware that the population defined by the weights for *each implicate* (see above) is 128.6 million households: the sum of each of the weights over all sample cases and imputation replicates is equal to five times the number of households in the sample universe.

Overview paragraph

Estimand paragraph

Results paragraph

Why it matters paragraph

Telegraphing paragraph: The remainder of this paper is structured as follows. Section 2....

2 Data

2.1 Overview

We use the statistical programming language R (R Core Team 2023).... Our data (Toronto Shelter & Support Services 2024).... Following Alexander (2023), we consider...

Overview text

2.2 Measurement

Some paragraphs about how we go from a phenomena in the world to an entry in the dataset.

2.3 Outcome variables

Add graphs, tables and text. Use sub-sub-headings for each outcome variable or update the subheading to be singular.

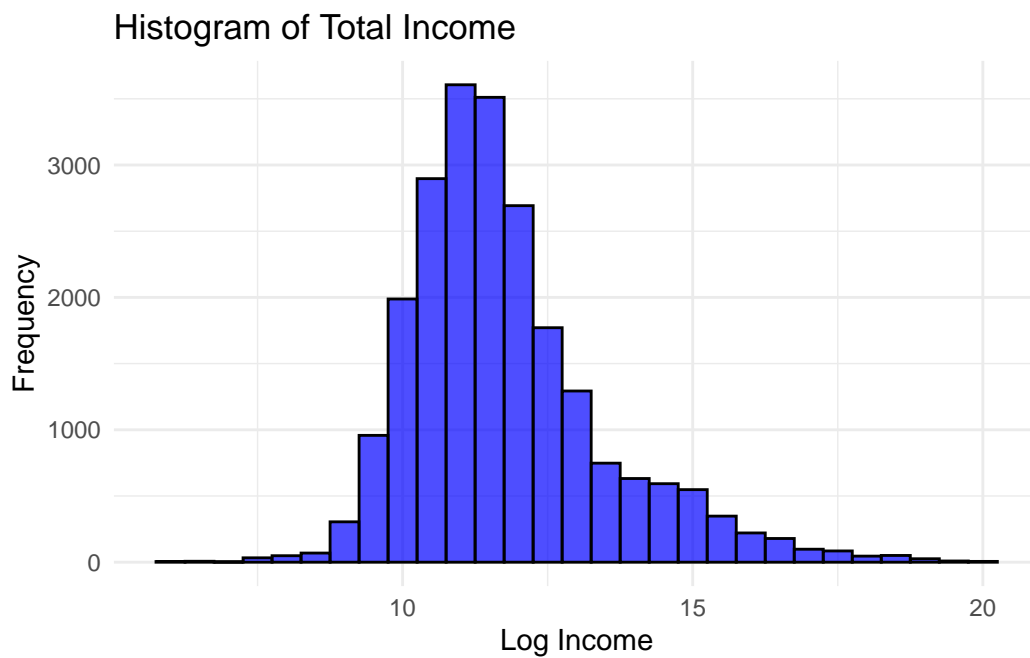Some of our data is of penguins (), from .

Figure 1: Histogram of Total Income
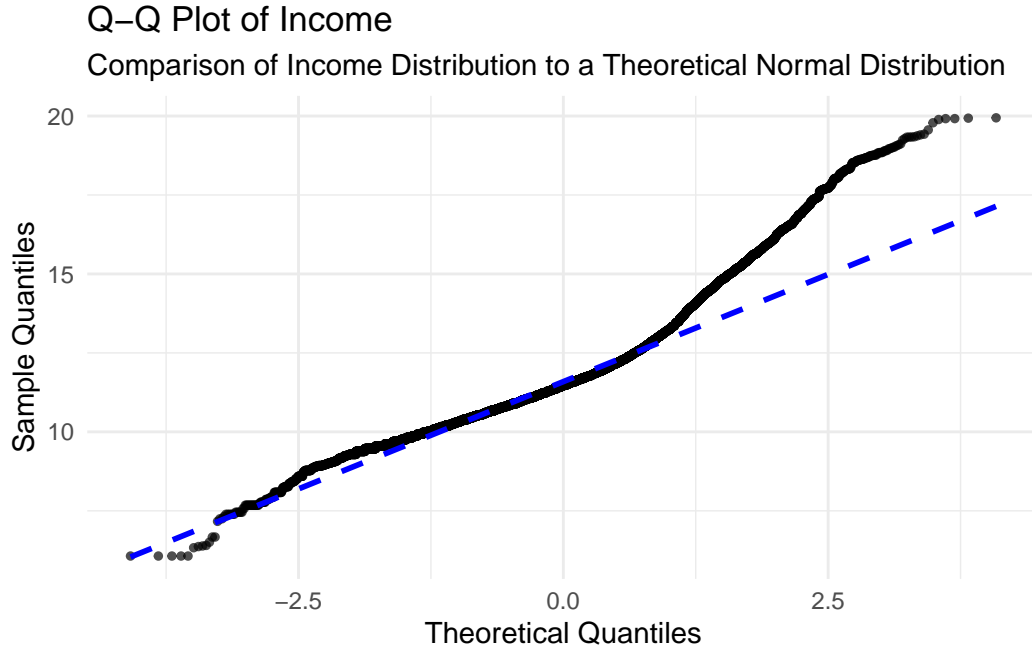
Talk way more about it.

Q–Q Plot of Income
Comparison of Income Distribution to a Theoretical Normal Distribution

Figure 2: Relationship between wing length and width

2.4 Predictor variables

Add graphs, tables and text.

Use sub-sub-headings for each outcome variable and feel free to combine a few into one if they go together naturally.

3 Model

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.
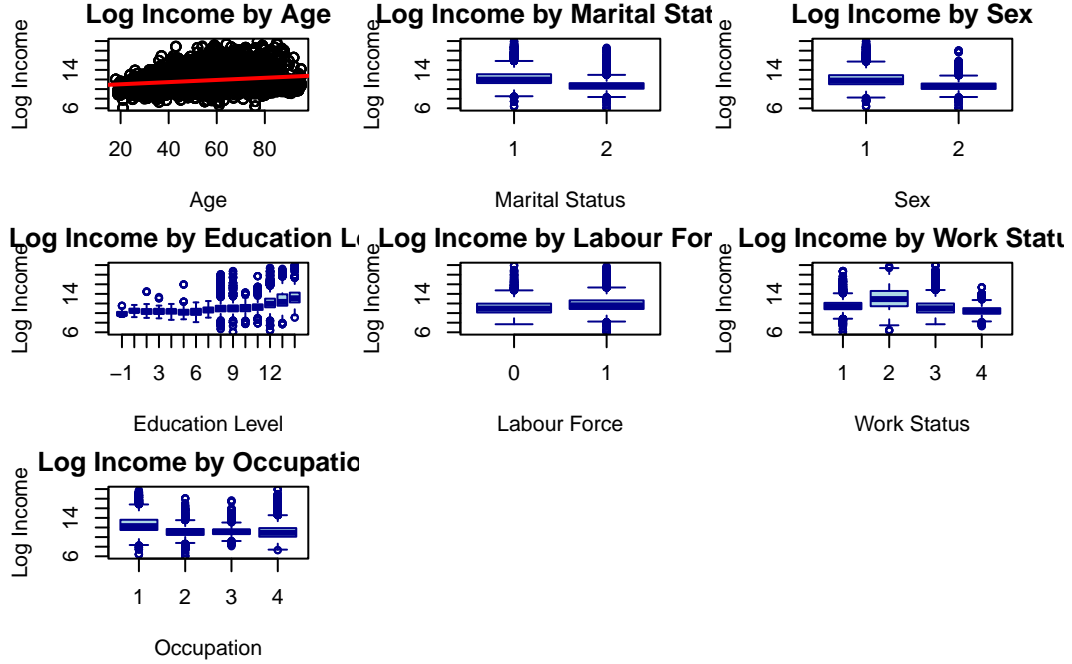
Figure 3: Relationship of Predictors variables with Total Income

$$y_i | \mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$

$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$

$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$

$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$

$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$

$$\sigma \sim \text{Exponential}(1) \tag{6}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance θ.

4 Results

Our results are summarized in Table **??**.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows...

In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows...

Examining how the model fits, and is affected by, the data

B.2 Diagnostics

?@fig-stanareyouokay-1 is a trace plot. It shows... This suggests...

?@fig-stanareyouokay-2 is a Rhat plot. It shows... This suggests...

Checking the convergence of the MCMC algorithm

References

Alexander, Rohan. 2023. *Telling Stories with Data.* Chapman; Hall/CRC. https://tellingstorieswithdata.com/.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "rstanarm: Bayesian applied regression modeling via Stan." https://mc-stan.org/rstanarm/.

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

Toronto Shelter & Support Services. 2024. *Deaths of Shelter Residents.* https://open.toronto.ca/dataset/deaths-of-shelter-residents/.